EPIGENETIC CHANGES DURING ACUTE OXIDATIVE DAMAGE AND NON-SMALL CELL LUNG CANCER EPIGENETIC THERAPY

by

Wei Wang

A dissertation submitted to Johns Hopkins University in conformity with the requirements for the degree of Doctor of Philosophy

Baltimore, Maryland

February, 2014

Abstract

Cancers of virtually all types exhibit not only genetic abnormalities but also epigenetic ones which consist of DNA methylation and chromatin alterations. Among the key cancer specific changes involving interplay of these epigenetic parameters are simultaneous wide spread losses of DNA methylation and more focal gains at CpG rich, proximal promoter regions and/or CpG islands. The latter change can associate with abnormal gene silencing and loss of gene function. These DNA methylation changes arise early in cancer risk states and early phases of tumor initiation, and may be important drivers of these processes. One key risk state for neoplastic transformation is chronic inflammation and oxidative stress, including stress generated during metabolic events. By exposing cells to hydrogen peroxide (H_2O_2) to mimic acute increase in reactive oxygen species, we demonstrate that, in addition to the known function of this insult for inducing DNA mutagenesis, oxidative damage can induce important epigenetic alterations in multiple cell line models, including the above changes in DNA methylation.

Upon H_2O_2 treatment, DNA methyltransferase 1 (DNMT1), the main enzyme for maintaining DNA methylation on newly replicated DNA, and sirtuin 1 (SIRT1), a NAD^+-dependent class III histone deacetylase, become more tightly bound to chromatin. SIRT1, a key player in epigenetic gene silencing and DNA damage repair, is involved in a transformation specific polycomb complex, polycomb repressive complex 4 (PRC4). Other components of PRC4 include enhancer of zeste homolog 2 (EZH2), which catalyzes the repressive histone modification, trimethylation of lysine 27 of histone H3. DNMT1, DNMT3B, one of the *de novo* DNA methyltransferases, and components of

PRC4 coexist in a large silencing complex(es). With H_2O_2 treatment, the size and abundance of this complex(es) increase. Co-immunofluorescence experiments indicate that this complex(es) resides, at least in part, at DNA damage foci. Through further genome-wide and local chromatin immunoprecipitation analyses, we show that constituents of this oxidative damage induced complex(es) concurrently translocate from transcriptionally inactive to active chromosome regions and from non-GC-rich to GC-rich areas, including promoter CpG islands, within 30 minutes of H_2O_2 exposure.

Accompanying the above re-localization, several functional consequences are observed. First, after treatment, there is reduction in the active histone marks, H3K4Me3 and H4K16Ac, and enrichment of the repressive H3K27Me3 mark at regions most targeted by members of the complex(es), for example, the active CpG island-containing gene promoters; whereas opposing trends are observed at low expression non-CpG island gene promoters. Second, there is reduction in the nascent transcription levels of the active CpG island-containing genes, *MYC*, *ACTB*, *TIMP3*, and *MLH1* within 30 minutes of treatment. In contrast, the repressive non-CpG island-containing genes, *NANOG*, *HBD*, and *IL8*, either gain or have no change in their nascent transcription levels. Third, it is found that genes targeted by the complex(es) often develop cancer-specific, abnormal, CpG island DNA hypermethylation during tumorigenesis. Indeed, several genes, such as *MLH1*, *SFRP5*, and *SFRP4*, show increased promoter CpG island DNA methylation, as determined by bisulfite sequencing after H_2O_2 treatment.

All of the above data provide a possible mechanism for abnormal changes in transcription and histone modifications often observed in cancer, and helps to explain the conundrum in the abnormalities of DNA methylation in cancer, namely, why cancer cells

simultaneously harbor both widespread chromosomal loss of DNA methylation and increased DNA methylation in CpG islands of gene promoters. More excitingly, some of the above observations were also validated by an *in vivo* colitis model, further strengthening the physiologic significance of our findings in cancer risk states, such as chronic inflammation.

A second series of studies were focused on the fact that aberrant epigenetic changes are more easily reversed by pharmacological interventions as compared to genetic mutations. Thus, increasing efforts are being exerted to develop preclinical and clinical studies of epigenetic drugs. Among these, the most studied are various DNA demethylating agents, such as 5-aza-cytidine (5-Aza-CR), and histone deacetylase inhibitors, such as Entinostat. A recent clinical trial with low dose 5-Aza-CR plus Entinostat has shown true therapeutic promise in patients with heavily pre-treated non-small cell lung cancer (NSCLC). To help understand this, we matched genome-wide DNA methylation and gene expression responses to 5-Aza-CR in eight NSCLC cell lines with basal levels of these parameters in the Cancer Genome Atlas (TCGA) project.

The response signature for the cell line with the most robust anti-tumor response, based on a *ex vivo* model, clusters with a small group of TCGA lung squamous tumors, and couples low expression and DNA hypermethylation of *RASSF1* with gene expression changes suggesting high cell cycle entry and stem cell pathway activity and low apoptotic capacity by pathway analyses. This suggests effective biomarker strategies to identify the best candidate patients for response to epigenetic therapy.

Meanwhile, the general response signature for all the eight cell lines, clusters with larger groups of TCGA tumors, and points to a reversal of a complex tumor immune

evasion signature, closely tied to interferon response pathways as analyzed by multiple pathway programs and gene set enrichment analysis (GSEA). A clue to the role of the drug in this process is that in some of the NSCLC lines, 5-Aza-CR induces the up-regulation of IRF7, a transcription factor for the interferon pathway, whose CpG island DNA hypermethylation and low expression tracks the immune evasion signature. FACS analysis also shows increased cell surface presence of PD-L1, an essential target for therapy to inhibit immune tolerance, by 5-Aza-CR treatment. These findings all suggest that epigenetic therapy may provide subsequent sensitization to therapy targeting immune tolerance checkpoints. In fact, in a trial of such immunotherapy, all 5 patients who happened to receive such therapy subsequent to 5-Aza-CR treatment have shown treatment benefit lasting at least 24 weeks, and three of the five patients have experienced major responses by RECIST criteria to subsequent immunotherapy, all of which are ongoing from 18 to 24+ months. We suggest this kind of pre-clinical, genomics based, approach could help efficient identification of patients who will receive maximum benefit from a breaking immune tolerance therapy given after epigenetic therapy. A larger trial has just started to test this hypothesis and many parameters of our above pre-clinical studies will be measured in patient tumor and blood samples to help develop our biomarker hypotheses.

Advisor: Dr. Stephen B. Baylin

Readers: Dr. Stephen B. Baylin and Dr. Kirby D. Smith

Acknowledgements

Numerous people have given me tremendous help and support throughout my years in graduate school. I would not be able to move this far without them. My sincere appreciation and thanks to all of them.

Foremost, I would like to express my deepest gratitude to my advisor, Dr. Stephen Baylin, for his continuous support and guidance throughout my Ph.D. study and research. Steve is an amazing scientist and great mentor. I am always inspired by his immense knowledge and by his dedication and passion to science. I am thankful to him for bringing me to the world of epigenetics, for teaching me how to conduct successful scientific research, and for encouraging me when work didn't go quite smoothly. I am especially thankful to him for his full support when I decided to pursue research further on the bioinformatics side.

Besides my advisor, I would also like to thank the rest of my thesis committee: Dr. Robert Casero, Dr. William Nelson, and Dr. Jef Boeke, for their insightful scientific input and tremendous support. In addition, my sincere gratitude goes to Dr. Kirby Smith, the second reader of my thesis dissertation, for his generous help during the writing of this thesis.

I am extremely fortunate to work with a group of fantastic colleagues who helped me not only in science but also shared with my happiness and struggles throughout my graduate study. Kevin Pruitt introduced the basics of epigenetics to me during my rotation. Kelly McGarvey taught me how to do ChIP-on-chip. Special thanks to Heather O'Hagan, who is a fantastic scientist and collaborated with me on the oxidative damage

story of my thesis. Heather is smart, knowledgeable and hardworking. It was an extremely enjoyable experience working with her, and I learned a lot from her. John Wrangle, who collaborated with me on the lung cancer story, is another amazing colleague and friend. He is full of great ideas. He broadened my view on both the clinic aspect as well as the scientific aspect of research, and we supported each other through the struggle of the paper. Hariharan Easwaran, our bioinformatic core in the lab is a very smart scientist. I enjoyed all the discussions I had with him. Yi Cai is another person I am hugely indebted to. He is kind of another mentor unofficially on both my research and my life. We shared a lot of intriguing discussion on science. He also taught me many experimental techniques and was there to help me unconditionally whenever I needed. I also owe a lot of thanks to many others in the group: Helai P. Mohammad, Eriko Clements, Christin Hanigan, Hsing-Chen Tsai, Subhojit Sen, Christina DeStefano Shields, Alexander Koch, Ray-Whay Chiu Yen, Tina Largent, etc. We shared a lot of stimulating discussions in science and a lot more laughter and tears in life. These are all priceless memories I will keep forever. I would also like to thank Kathy Bender for her excellent work on keeping the lab functional and making our life easier.

I really appreciate the support I have received from the Human Genetics Training Program. I want to especially thank Sandy Muscelli, our program coordinator, for taking care of us throughout our graduate education. Besides, I am very grateful to people from the Biostatistics Department from School of Public Health, Dr. Hongkai Ji, Dr. Brian Caffo, Mary Joy Argo, etc., for providing me the chance to get thorough training in biostatistics and helping me in getting my concurrent MHS degree. I would also like to thank my dearest friends in Hopkins: Yangfan Liu, Yin Liu, and Yingxin Zhang, etc.

Thank you for your generous support and thank you for all the fun you have brought to my life.

Last but not the latest, a special thanks to my family for their constant love, support, and belief in me. Thank my parents for bringing me up and supporting me through all the steps of my life. Words cannot express how grateful I am for all of the sacrifices that you've made on my behalf. And at the end, my deepest appreciation to my beloved husband, Zhuhao Wu, who has accompanied me through all my ups and downs, and is always my strongest support all the time.

Table of Contents

List of Tables

List of Figures

Intended to be blank

Chapter 1

Introduction

1.1 Epigenetics

The prefix “epi-”, derived from Greek, means “in addition to”. Thus, it is easy to understand “epigenetics” as referring to phenotypes associated with mechanisms other than those inherited genetically. Since the invention of this word by British scientist C. H. Waddington in 1942 (Waddington, 2012), the definition of epigenetics has been modified by many others, including Robin Holliday, Arthur Riggs, and Adrian Bird et al. A recent consensus definition of epigenetics was achieved at the 2008 Cold Spring Harbor meeting, as the “stably heritable phenotype resulting from changes in a chromosome without alterations in the DNA sequence” (Berger et al., 2009).

Common epigenetic mechanisms of gene expression regulation include, but are not limited to, DNA methylation, chromatin remodeling, and histone modifications. Coordinately, these mechanisms define the epigenetic landscape of cells, from the time of embryogenesis through adult cell renewal, which establish their different phenotypes despite their having identical genetic information. The heritably of the epigenetic control can then be maintained through generations of cell division.

Abnormality in the control of any of these epigenetic mechanisms can cause severe consequences, which can manifest during development and in diseases such as cancer. It has been shown that epigenetic aberrations at both individual gene levels and on a genome-wide scale can underlie these abnormalities. Identifying these aberrations and fully understanding these epigenetic mechanisms will significantly aid in management of diseases such as for the detection and treatment of cancer.

1.1.1 DNA methylation

Biological functions

DNA methylation is found in the genomes of organisms ranging from prokaryotes to eukaryotes. In prokaryotes, DNA methylation occurs on both cytosine and adenine bases and constitutes part of the host restriction system (Wilson and Murray, 1991). In multicellular eukaryotes, however, the majority of DNA methylation is confined to cytosine bases and is associated with a condensed chromatin state and repression of gene expression (Bird, 2002). This methylation, for cytosine, is catalyzed by a family of enzymes named DNA methyltransferases (DNMTs) and involves the transfer of a methyl group from a S-adenosyl methionine donor, to the 5' position of the pyrimidine ring of cytosine (Cheng and Blumenthal, 2008). In humans, the methylation of cytosine usually occurs at CpG dinucleotides (Bird, 1980). It is well established that this methylation in gene proximal promoter regions is typically associated with transcriptional repression, whereas the methylation of gene bodies correlates with transcriptional activity, probably secondary to facilitation of transcriptional elongation (Varley et al., 2013). Through such dynamics, 5-methylcytosine modulates multiple biological processes including embryonic development, gene imprinting, often gene silencing but also gene activation in some circumstances as mentioned, and the structural integrity of chromosomes (Bird, 2002).

DNA methylation is not distributed evenly in the mammalian genome. The CpG dinucleotide is under-represented in the human genome due to, over evolution, the high mutation rate of 5-methylcytosine to thymine (Bird, 1986). Thus, the remaining CpG sites are heavily DNA methylated in the human genome (Bird, 1986). The exception is so

called “CpG islands”, defined as GC-rich regions (GC content ≥ 55%) of 500 base pairs (bp) or longer that possess high densities of CpG dinucleotides (observed to expected CpG ratio ≥ 0.65) (Takai and Jones, 2002). Computational genome analysis predicts the existence of ~29,000 CpG islands in the human genome associated with ~60% of gene promoters (Antequera and Bird, 1993; Venter et al., 2001). The majority of these promoter CpG islands are unmethylated at all developmental stages in all tissue types (Antequera and Bird, 1993). A small proportion of these CpG islands become methylated during development and are associated with stable gene silencing. Such silencing, on an individual allele associated basis, is developmentally programmed for allelic retention of CpG-island methylation especially for genomic imprinting and X chromosome inactivation (Bird, 1993). *De novo* DNA methylation occurs in germ cells or the early embryo (Jaenisch et al., 1982), suggesting that it is particularly active at these stages. In addition, *de novo* DNA methylation can also occur in the aging process of adult somatic cells. A significant proportion of human CpG islands are prone to progressive methylation during aging (Issa, 2000), or in abnormal cells such as cancer cells (Baylin and Herman, 2000; Plass et al., 2013; Shen and Laird, 2013).

Although DNA methylation predominately occurs at CpG dinucleotides in the human genome, there have been some reports that non-CpG cytosine methylation exists in embryonic and pluripotent cells. First, high-throughput sequencing of bisulphite treated DNA showed that in human embryonic stem (ES) cells 25% of 5-methylcytosine was in a non-CpG context, whereas in fetal lung fibroblast cell line IMR90 only 0.02% of methylcytosine was in a non-CpG context, despite overall amounts of CpG methylation being similar (Lister et al., 2009). Cytosine methylation in non-CpG contexts shows

enrichment in gene bodies, with the level of methylation positively correlating with gene expression levels (Lister et al., 2009). The ENCODE project found that non-CpG methylation also exists in somatic cells, especially in the brain (Varley et al., 2013). The biological function and molecular mechanism of non-CpG methylation are not clear and are worth future investigation.

Multiple lines of evidence support the idea that the major biological function of DNA methylation is to regulate chromatin structure. The most compelling evidence for this idea comes from DNA transfection studies showing that unmethylated DNA substrates are packaged into an open chromatin structure following their integration into the genome and are sensitive to DNase I treatment, whereas the same DNA remains completely resistant to DNase I treatment if it is methylated (Keshet et al., 1986). Currently, the molecular mechanism by which DNA methylation regulates chromatin packaging is unclear. Early experiments indicated that most methylated CpG dinucleotides are concentrated within the central core of nucleosomes, as opposed to inter-nucleosomal regions, suggesting that the positioning of methylated CpG dinucleotides may have an intrinsic effect on where nucleosomes reside on the DNA (Razin and Cedar, 1977; Solage and Cedar, 1978), and these results have recently been confirmed by sophisticated genome-wide analysis (Chodavarapu et al., 2010). Another way DNA methylation mediates chromatin structure is through methyl-binding proteins, such as MBD2. These proteins specifically recognize methylated CpG dinucleotide and remodel local chromatin structure through recruiting histone deacetylases HDAC1/2 or histone demethylase LSD1 (Lai and Wade, 2011). Alternatively, DNA methylation may block the binding of chromatin proteins such as chromatin boundary factor CTCF or Cfp1, which are known

to exclusively bind unmethylated CpG islands (Bell and Felsenfeld, 2000; Thomson et al., 2010). These results provide important insights into the mechanism of epigenetic gene silencing mediated by DNA methylation.

Mechanisms for establishing

DNA methylation is a post-synthetic modification (Burdon and Adams, 1969; Scarano et al., 1965). As previously mentioned, this process is catalyzed by a family of enzymes known as DNMTs. In mammals, four DNMTs have been identified, with different roles in the establishment and maintenance of DNA methylation (Bird, 2002).

DNMT1 has the ability to copy the methylation pattern from one strand to another and has led to DNMT1 being known as the maintenance methyltransferase. It was thought that DNMT1 allows the faithful replication of the 5-methylcytosine marks onto the newly synthesized DNA strand through its association with the master regulator protein of DNA replication, PCNA (Chuang et al., 1997). However, recent work revealed that UHRF1 is the major protein which targets DNMT1 to hemi-methylated DNA (Bostick et al., 2007; Sharif et al., 2007). In addition, DNA methylation is well maintained in a DNMT1 hypomorph of human colon cancer cell line HCT116 which express a truncated DNMT1 isoform without its PCNA binding motif (Egger et al., 2006). The importance of DNMT1 protein can be inferred from the lethality of the Dnmt1 knockout mouse (Li et al., 1992) and the lethal effect of DNMT1 complete knockout in human colon cancer cell line HCT116 (Chen et al., 2007). Both the DNMT1 knockout mouse and ES cells show approximately 70% reduction in 5-methylcytosine. Although DNMT1 knockout ES cells are viable, the DNMT1 knockout mouse is embryonic lethal and embryos are stunted and die at midgestation (Li et al., 1992). DNMT1 has been

shown to interact with histone methyltransferase G9a and histone deacetylases HDAC1/2, raising the possibility that DNMT1 functions, in addition to catalyzing DNA methylation, as a scaffold protein to recruit other epigenetic modifier proteins to modulate local chromatin states (Robertson et al., 2000; Rountree et al., 2000).

Through searching for human proteins that were homologous to bacterial DNA methyltransferases, *de novo* DNA methyltransferases DNMT3A and 3B were discovered (Okano et al., 1998). These two proteins were most highly expressed in ES cells and *in vitro* DNA methylation studies showed that recombinant DNMT3A and DNMT3B proteins act on non-methylated DNA and had no preference for hemi-methylated DNA (Okano et al., 1998). Genetic knockouts of Dnmt3a and 3b in mouse ES cells indicated that, in the absence of both proteins, there was a loss of *de novo* DNA methylation, but single knockouts still maintained *de novo* DNA methylation activity, suggesting that DNMT3A and DNMT3B have redundant functions (Okano et al., 1999). Methylation analysis of the genomic DNA from single DNMT3 knockout ES cells revealed that loss of Dnmt3a resulted in hypomethylation of major satellite repeats and a loss of some maternal imprints, whereas loss of Dnmt3b led to hypomethylation of minor satellite repeats (Chen et al., 2003; Okano et al., 1999). This result suggested that DNMT3A and DNMT3B target different genomic regions.

Although DNMT1 is thought to be the maintenance DNMT, there is some evidence that DNMT3A and DNMT3B also participate in DNA methylation maintenance. Extended culture of Dnmt3a/3b double knockout ES cells led to a progressive loss of methylation to levels approaching that of Dnmt1 knockout ES cells (Chen et al., 2003). Overall, this suggests that both DNMT3A and 3B are required in concert with DNMT1

for the maintenance of DNA methylation, arguing against the simple model of "maintenance" and "*de novo*" methyltransferases (Bird, 2002). For this reason, Peter Jones suggested a new model in which all three DNMTs work in concert to maintain methylation patterns with DNMT3A and 3B helping to fill in methylation gaps left by DNMT1 (Jones and Liang, 2009).

Recently, hydroxymethylation at the C-5 position of cytosine in the context of CpG dinucleotides was observed in vertebrate genomes (Tahiliani et al., 2009). In contrast to 5-methylcytosine which is catalyzed by DNMTs, the conversion of 5-methylcytosine to 5-hydroxymethylcytosine is catalyzed by the TET (ten-eleven translocation proteins or methylcytosine dioxygenases) family of enzymes (Tahiliani et al., 2009). 5-hydroxymethylcytosine can be converted back to unmodified cytosine either passively or actively resulting in DNA demethylation (Pfeifer et al., 2013). This is currently the only known method for active DNA demethylation. However, final conversion of 5-hydroxymethylcytosine back to cytosine requires additional DNA repair mediated steps and these can act, also, in concert for conversion of 5-methylcytosine back to cytosine (Bhutani et al., 2011; Piccolo and Fisher, 2013). Recent results suggested that 5-hydroxymethylcytosine might also mediate transcriptional repression through the recruitment of methyl-binding protein MBD3 (Yildirim et al., 2011).

1.1.2 Chromatin remodeling

In eukaryotes, DNA is packaged into basic structure units, called nucleosomes. From a linear standpoint, each nucleosome core particle encompasses a segment of DNA of about 147bp wrapped around a histone octamer, consisting of pairs of each histone core

proteins, H2A, H2B, H3, and H4, or their variants (Lilley and Pardon, 1979; McGhee and Felsenfeld, 1980). Nucleosome core units are connected by linker DNA of various lengths and the linker histone protein H1 and then compacted into higher order chromatin structures (Felsenfeld and McGhee, 1986; Struhl and Segal, 2013). When this packing is tight and highly organized, the landscape is termed heterochromatin and is associated with a transcriptionally repression state (Fahrner and Baylin, 2003; Wutz, 2011). In contrast, euchromatin represents areas of much less tightly packed DNA with often nucleosome free regions termed as relatively open chromatin and this is often associated with active transcription (Bassett et al., 2009; Lamond and Earnshaw, 1998).

The dynamic balance between tight and more open chromatin packaging is important for epigenetic regulation of gene expression (Bassett et al., 2009). Chromatin dynamics involves interaction among histones, histone chaperone proteins, histone modification enzymes, and a group of chromatin remodeling complexes (remodelers) (Allis et al., 2007; Khorasanizadeh, 2004). Chromatin remodelers use the energy of ATP hydrolysis to change the packaging state of the nucleosome which is the primary structure unit of chromatin and to regulate where it is located (Clapier and Cairns, 2009). Remodelers cooperate with other chromatin factors to regulate the packaging and unpackaging of genomic regulatory DNA elements including enhancers, promoters, replication origins that are essential for biological processes such as gene transcription, DNA replication and DNA repair (Altmeyer and Lukas, 2013; Fadloun et al., 2013; Soria et al., 2012). Therefore, chromatin structure becomes a key platform for regulation of gene expression. Four major types of chromatin remodeler families and their biological functions in gene repression, DNA damage and DNA repair are listed below.

- SWI family remodelers**.** The SWI/SNF (switching defective/sucrose nonfermenting) family of remodelers was first characterized in *Saccharomyces cerevisiae* and contains 8 to 14 protein subunits. This family of remodelers has many activities including sliding and ejecting nucleosomes at many genomic loci (Mohrmann and Verrijzer, 2005).
- ISWI family remodelers. The ISWI (imitation switch) family of remodelers contains 2 to 4 protein subunits. dNURF, dCHRAC, and dACF complexes were originally identified in *Drosophila melanogaster*, with hWICH or hNoRC characterized in human subsequently. Many ISWI family complexes (ACF, CHRAC) optimize nucleosome spacing to promote chromatin assembly and repress transcription. However, some ISWI complexes (NURF) can randomize spacing, which can activate transcription (Corona and Tamkun, 2004).
- CHD family remodelers. The CHD (chromodomain, helicase, DNA binding) family of remodelers is composed of 1 to 10 protein subunits and was initially purified from *Xenopus laevis* (Marfella and Imbalzano, 2007). Certain CHD remodelers slide or eject nucleosomes to promote transcription. However, others have repressive roles, including the vertebrate Mi-2/NuRD (nucleosome remodeling and deacetylase) complex, which contains two histone deacetylases (HDAC1/2) and two methyl CpG-binding domain proteins, MBD2 and MBD3 (Denslow and Wade, 2007).
- INO80 family remodelers. The INO80 (inositol requiring 80) family of remodelers contains more than 10 protein subunits and was first characterized from *Saccharomyces cerevisiae* (Bao and Shen, 2007). Its higher orthologs include hINO80 and p400, which also contains histone acetyltransferase activity. INO80

family of remodelers has multiple biological functions including promoting transcriptional activation and DNA repair (Min et al., 2013).

The general mechanism for chromatin remodelers to repress transcription is to help genomic regions assemble into nucleosome arrays, restricting the access of DNA-binding transcription factors (Hargreaves and Crabtree, 2011; Narlikar et al., 2013). Furthermore, remodelers also cooperate with HDAC to repress transcription repression (Marks et al., 2001). For example, the chromatin remodeling complex NuRD contains histone deacetylases HDAC1 and HDAC2, and a methyl-DNA binding protein (MBD2) for recognizing DNA methylation, strongly supporting its role in the maintenance of gene silencing in methylated genomic regions (Clapier and Cairns, 2009). Another good example is the NoRC complex, which mediates rDNA silencing through its recruitment to the RNAPI promoters by TTF-I (Strohner et al., 2004); then NoRC remodels the promoter-bound nucleosome to a position unfavorable for transcription and recruits HDAC and DNMTs in an H4K16ac-dependent manner (Li et al., 2006). Finally, the *S. pombe* SHREC complex is important for the assembly of silent pericentromeric heterochromatin (Sugiyama et al., 2007).

1.1.3 Histone modifications

The N-terminal tails of histone proteins are freely extended from the core structure and are actively subjected to covariant modifications, such as methylation, acetylation, phosphorylation, ubiquitination, and sumoylation (Berger, 2002; Jenuwein and Allis, 2001). The type and position of each histone modification, and the cross-talk between different modifications can uniquely alter protein-DNA and protein-protein interactions

that affect chromatin configuration and recruitment of other chromatin-binding proteins, thus impacting gene expression in either an active or repressive manner (Strahl and Allis, 2000; Zhang and Reinberg, 2001). Therefore, the modifications of histones establish another layer of epigenetic code, the histone code (Jenuwein and Allis, 2001; Wang et al., 2004).

Some of the most commonly studied histone modifications include: the trimethylation of lysine 4 on histone H3 (H3K4Me3), acetylation of lysine 9 and lysine 14 on histone H3 （H3K9Ac and H3K14Ac), and acetylation of lysine 16 on histone H4 （H4K16Ac), are specially enriched at actively transcribed gene promoters (Eissenberg and Shilatifard, 2010; Graff and Tsai, 2013; Zentner and Henikoff, 2013). The trimethylation of lysine 36 on histone H3 (H3K36Me3) is usually found at actively transcribed gene bodies (Wagner and Carpenter, 2012). In contrast, the trimethylation of lysine 9 on histone H3 (H3K9me3) is highly correlated with constitutive heterochromatin and the trimethylation of lysine 27 on histone H3 (H3K27me3) is frequently observed at transcriptionally repressed gene promoters and is heavily enriched at facultative heterochromatin (Bracken and Helin, 2009; Shinkai and Tachibana, 2011). Sometimes, proximal promoter chromatin domains can be marked by both the active H3K4me3 mark and the repressive H3K27me3 mark, forming so-called bivalent chromatin (Bernstein et al., 2006; Ohm et al., 2007). Such chromatin domains are especially prominent in ES cells (Johnson et al., 2012; Marks et al., 2012). Genes which are so marked, constituting ~ 2500 in ES cells, are often poised for activation of transcription, giving the possibility to rapidly alter to fully active or silenced states with differentiation of ES cells into distinct lineages as necessary (Bernstein et al., 2006; Spivakov and Fisher, 2007). These genes that are bivalently

marked in ES cells are often the targets of polycomb repressive complexes in adult stem cells and have a higher tendency to be aberrantly silenced by DNA hypermethylation in cancer (Easwaran et al., 2012; Ohm et al., 2007; Widschwendter et al., 2007).

Numerous enzymes are involved in applying and removing of specific type of histone modifications. The acetylation and deacetylation of lysine are quite dynamic and carried out by two large families of enzymes, histone acetyltransferases (HATs) and histone deacetylases (HDACs), respectively (Bannister and Kouzarides, 2011; Zentner and Henikoff, 2013). One the other hand, enzymes involved in the regulation of methylation of lysine, histone lysine methyltransferases (HKMT) and histone demethylases (HDMs), have greater substrate specificity (Bannister and Kouzarides, 2011; Zentner and Henikoff, 2013). For example, SETD7 specifically monomethylates lysine 4 on histone H3, and KDM4C demethylates H3K9Me3 but not H3K9Me or H3K9Me2 (Berry and Janknecht, 2013; Wagner and Jung, 2012).

In addition, there are specific histone modifications which accompany DNA damage and repair processes (Price and D'Andrea, 2013). When DNA damages occur in the genome, a rapid cellular response is the phosphorylation of serine 129 of H2A in yeast or serine 139 of H2A.X in vertebrates (Xiao et al., 2009). This histone modification helps recruit DNA repair factors as well as chromatin remodelers of the SWI/SNF and INO80 families which facilitate the access of DNA repair enzymes to the DNA damage sites (Lukas et al., 2011; Xiao et al., 2009). Therefore, the phosphorylation of H2A/H2A.X provides a good system for the rapid recognition of DNA damage sites followed with efficient repair.

1.2 Epigenetics in Cancer

In addition to genetic mutations and chromosome translocations, aberrant epigenetic regulation is often observed in cancer and is a significant contributor to tumorigenesis (Jones and Baylin, 2007). Aberrant activation of oncogenes could be a direct trigger for tumorigenesis, whereas epigenetic silencing of commonly expressed tumor suppressor genes can provide for loss of function of these genes (Baylin and Bestor, 2002). These abnormalities often occur in the setting of one of the key features of cancer cells, abnormal DNA-methylation patterns which include gain of methylation (hypermethylation) and loss of methylation (hypomethylation) (Feinberg, 2007). It is well established that cancer progression associates with a decrease of overall DNA methylation and an increase in the methylation of CpG islands located in the promoter regions of many tumor suppressor genes (Jones and Baylin, 2007). Recently, the ENCODE project did a base-resolution mapping of methylcytosine across 82 cell lines including both normal and cancer cells (Varley et al., 2013). This study discovered that there are 66,570 CpGs hypermethylated in cancer cells, including cell lines derived from breast, prostate, lung, ovarian, endometrial, liver, and pancreatic cancer, as well as neuroblastoma and several leukemias. Among these hypermethylated CpGs, 48,787 (73%) reside in CpG islands (Varley et al., 2013). The observation of cancer-specific CpG islands hypermethylation across the genome in multiple cancer cell lines is consistent with reports of a CpG island methylator phenotype (CIMP) that was first described in colorectal cancer (Toyota et al., 1999) and later documented in many other cancer types (Chen et al., 2012; Liu et al., 2011; Teodoridis et al., 2008; Turcan et al., 2012). An

additional 7377 (11%) non-island CpGs are significantly hypermethylated in cancer, and these reside in promoters and bodies of genes encoding proteins with sequence-specific DNA binding transcription factor activity (Varley et al., 2013). DNA hypermethylation in both the promoter and gene body of these transcription factor genes indicates dysregulation of methylation at these genes in various cancers. Chapter 2 of this thesis focuses on revealing the underline mechanisms of aberrant epigenetic silencing during tumorigenesis in a special background of oxidative stress which is common many cancer risk states, such as chronic inflammation.

1.3 Epigenetic Therapy

Epigenetic silencing of tumor suppressor genes is a general phenomenon during tumorigenesis in virtually all types of cancer (Esteller et al., 2001; Herman and Baylin, 2003; Jones and Baylin, 2007; Santini et al., 2001). Unlike genetic mutations, epigenetic alterations are pharmacologically reversible, which makes them attractive targets in cancer management (Issa and Kantarjian, 2009). Many compounds have been developed to specifically target proteins that control DNA methylation and histone modifications. Among them, two important DNA demethylating agents, 5-aza-cytidine and 5-aza-2'-deoxycytidine, have been of great interest.

5-aza-cytidine and 5-aza-2'-deoxycytidine were first synthesized in the 1960s as cytidine analogues. Once they are transported into the cells, they can be incorporated into DNA (5-aza-cytidine and 5-aza-2'-deoxycytidine) and RNA (5-aza-cytidine) after multiple phosphorylation steps (Oki et al., 2007). These drugs can form a covalent

complex with DNMTs, therefore inhibit the enzymatic activities of DNMTs and lead to DNA demethylation (Jones and Taylor, 1980; Jones and Taylor, 1981). Multiple reports indicated that 5-aza-cytidine and 5-aza-2'-deoxycytidine were able to induce durable complete remission in hematological malignancies, especially in myelodysplastic syndrome (MDS) at low doses (Cashen et al., 2010; Issa et al., 2004; Kantarjian et al., 2003; Wijermans et al., 2000). These results were further confirmed by larger clinical trials (Kantarjian et al., 2006; Silverman et al., 2002). Based on their promising clinical efficacy, 5-aza-cytidine and 5-aza-2'-deoxycytidine were approved by U.S. Food and Drug Administration (FDA) in 2004 and 2006 respectively for the treatment of MDS and have now become a standard care option for patients with high risk MDS.

The achievement of high clinical efficacy and low toxicity with low dose regimens opened more opportunities for the drugs. Mild toxicity profiles of the drugs at low doses enable clinicians to explore the possibility of combination therapies with other epigenetic-modifying agents, such as HDAC inhibitors, or with other chemotherapeutic agents or immunotherapeutic agents in both hematological malignancies and solid tumors (de Vos and van Overveld, 2005; Gore, 2005; Groselj et al., 2013). Numerous clinical trials are currently under-going and give us a promise for new therapy concepts aimed to markedly improve cancer management. Chapter 3 of this thesis focuses on pre-clinical work aimed at understanding the molecular basis of anti-tumor responses for the DNA demethylating agents discussed above, with an aim to develop biomarkers which predict efficacy.

Chapter 2

Oxidative damage targets complexes containing DNA methyltransferases, SIRT1 and polycomb members to promoter CpG islands

Reprint from: Oxidative damage targets complexes containing DNA methyltransferases, SIRT1, and polycomb members to promoter CpG Islands

Cancer Cell. 2011 Nov 15;20(5):606-19.

2.1 Summary

Cancer cells simultaneously harbor global losses and gains in DNA methylation. We demonstrate that inducing cellular oxidative stress by treatment with hydrogen peroxide, recruits DNA methyltransferase 1 (DNMT1) to damaged chromatin. DNMT1 becomes part of a complex(es) containing DNMT3B and members of Polycomb Repressive Complex 4. Hydrogen peroxide treatment causes translocalization of these proteins from non-GC-rich to GC-rich areas. Key components are similarly enriched at gene promoters in an *in vivo* colitis model. While high expression genes enriched for members of the complex have histone mark and nascent transcription changes, CpG island-containing low expression genes gain promoter DNA methylation. Thus, oxidative damage induces formation and localization of a silencing complex that may explain cancer-specific aberrant DNA methylation and transcriptional silencing.

2.2 Significance

Tumors have aberrant gains and losses in DNA methylation, although the mechanisms establishing these changes are not well understood. Here we demonstrate that oxidative damage induces the formation of a large silencing complex(es) containing DNA methyltransferases and constituents of the polycomb complex, PRC4, including SIRT1. PRC4 is found uniquely in cancer and embryonic and adult stem cells. Key constituents of the damage-induced complex are recruited from transcriptionally poor regions of the genome to GC-rich areas, including promoter CpG islands. Such translocalization causes changes in histone marks, transcription, and DNA methylation. We suggest that this re-localization may be a mechanism by which oxidative damage can be responsible for both promoter CpG island specific hypermethylation and global hypomethylation seen in cancer.

2.3 Highlights

- DNMT1 becomes more tightly bound to chromatin after oxidative damage.
- Oxidative damage induces formation of a complex containing DNMT1, DNMT3B and PRC4.
- DNMT-PRC4 enrichment at CpG islands may explain aberrant gene silencing in cancer.
- Promoters enriched for these proteins have histone mark and DNA methylation changes

2.4 Introduction

Elevated levels of reactive oxygen species (ROS) arising from alterations in cellular metabolism and inflammatory responses constitute a key risk state for increased cancer susceptibility (Federico et al., 2007). The major forms of oxidative DNA damage are nonbulky lesions such as 8-oxo-2'deoxyguanosine (8-oxo-dG) and thymine glycol that are repaired predominantly by base excision repair (BER) (Reardon et al., 1997).

The above DNA repair requires dynamic changes in surrounding chromatin including changes in nucleosome positioning and histone modifications. The best characterized chromatin alteration in DNA repair is the phosphorylation of the histone variant H2AX (γ-H2AX) by DNA damage response protein kinases (Rogakou et al., 1998). This modification helps stabilize the interaction of repair factors with the break sites, leading to further chromatin alterations. Histone acetylases and deacetylases also localize to sites of DNA damage to facilitate repair by increasing access of repair proteins to the break site, repressing transcription at sites of damage, restoring the local chromatin environment after repair is complete, and turning off the DNA damage response (Tamburini and Tyler, 2005). In this regard, (Sirtuin-1) SIRT1 is a NAD^+-dependent class III histone deacetylase that plays a role in gene silencing in cancer cells (Pruitt et al., 2006) and has been implicated in DNA damage repair in both yeast and mammalian cells. SIRT1 is recruited to sites of DNA damage and interacts with and deacetylates other proteins involved in the DNA damage (For review, see (Fan and Luo, 2010)). After DNA repair, DNA methylation also needs to be reestablished, possibly by the recruitment of the DNA methyltransferases (DNMTs) that catalyze CpG methylation, including

DNMT1 which plays a role in methylating newly replicated DNA (Leonhardt et al., 1992), and DNMT3A and DNMT3B which are mostly responsible for *de novo* DNA methylation (Okano et al., 1999).

The above epigenetic players have been linked to patterns of cancer-related gene transcriptional silencing, in association with promoter CpG island DNA hypermethylation. We, and others, have shown that a large fraction of the genes that undergo promoter CpG Island DNA hypermethylation in cancer are unmethylated in embryonic stem and progenitor cells and held in low/poised states of transcription by polycomb group (PcG) proteins (Ohm et al., 2007; Schlesinger et al., 2007; Widschwendter et al., 2007). Importantly, SIRT1 has been described as part of a transformation specific PcG complex, PRC4, which is found in embryonic and adult stem cells and cancer cells (Kuzmichev et al., 2005). In addition to SIRT1, the PRC4 complex contains the PcG proteins, Enhancer of Zeste protein-2 (EZH2) which catalyzes the trimethylation of lysine 27 of histone H3 and a specific isoform of EED (EED2) that is absent from previously identified PRC complexes. SIRT1 has also been shown to interact with DNMT1 (Espada et al., 2007).The DNMTs have been linked to PcG proteins in the context of epigenetic gene silencing. Both DNMT1 and DNMT3B interact with EZH2, which in turn facilitates the binding of the DNMTs to EZH2 target promoters (Vire et al., 2006).

In the present study, we investigate epigenetic alterations induced by the ROS, hydrogen peroxide (H_2O_2), and by inflammation in mouse tissue. We examine changes in the interaction and chromatin binding of the epigenetic proteins discussed above and the functional consequences of these changes. This work attempts to determine a mechanism

by which cancer risk states, such as chronic inflammation, can contribute to cancer-related abnormal gene silencing and shifts in DNA methylation.

2.5 Results

2.5.1 DNMT1 and SIRT1 Become Tightly Bound to Chromatin after H_2O_2 Treatment

Previously, we have demonstrated that SIRT1 and DNMT1 are rapidly recruited to an induced double strand break in an exogenous promoter CpG island construct (O'Hagan et al., 2008). In this regard, SIRT1, similar to other proteins involved in DNA repair, is known to become more tightly bound to chromatin after oxidative stress (Oberdoerffer et al., 2008). We now find, by examining resistance of the proteins to salt gradient extraction, that both SIRT1 and DNMT1 bind more tightly to chromatin in H_2O_2-treated human embryonic carcinoma cells (NCCIT) despite their unchanged whole cell levels. As evidence of this tightening, after H_2O_2 treatment, a portion of SIRT1 is redistributed from the cytoplasmic fraction to the soluble nuclear fraction and is present in all higher salt fractions (Figure 2.1A). Basally, as has been previously demonstrated, nuclear DNMT1 is loosely bound to the chromatin, being extracted by 0.3 and 0.45 M NaCl (Jeong et al., 2009). However, after H_2O_2 treatment, DNMT1 is also eluted in salt fractions of 0.6 M, 1.2 M, and 1.8 M NaCl (Figure 2.1A). HSP90 and LaminB serve as cytoplasmic and nuclear controls, respectively, for the extraction.

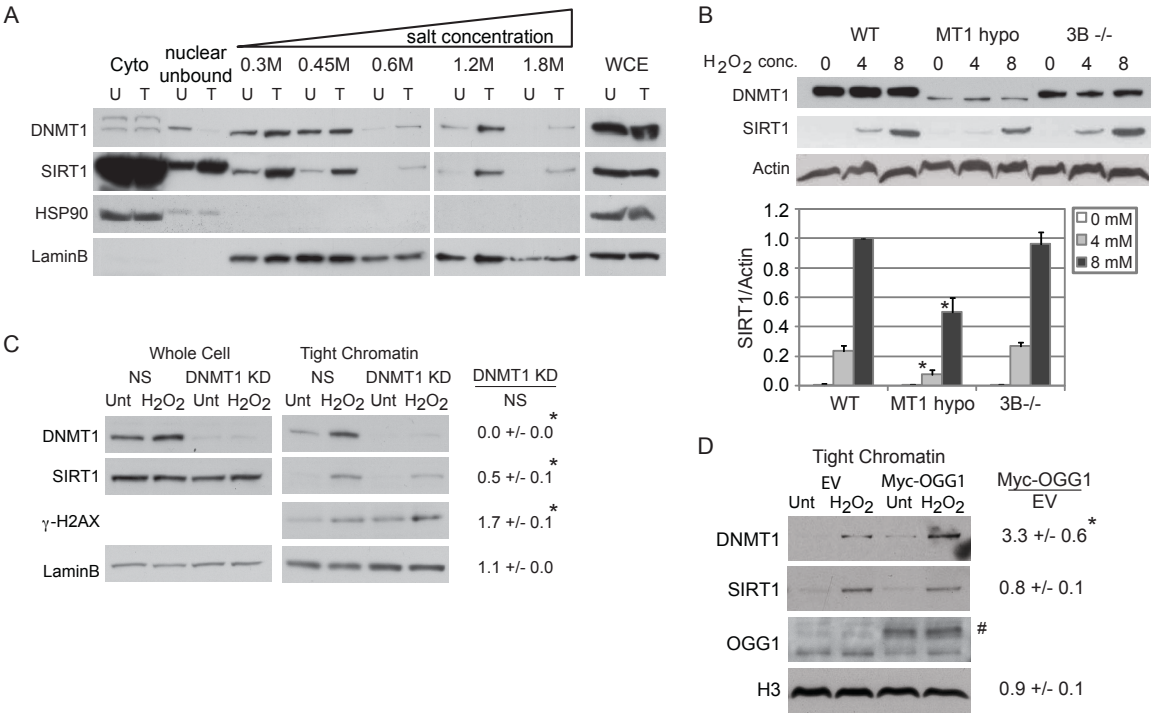


Figure 2.1. DNMT1 and SIRT1 become tightly bound to chromatin after treatment with H_2O_2.

(A) NCCIT cells were untreated (U) or treated with 1 mM H_2O_2 for 30 minutes (T). Cell pellets were extracted sequentially using cytoplasmic extraction buffer (Cyto), soluble nuclear buffer (nuclear unbound), and buffers with increasing NaCl concentration. Whole cell lysates were prepared separately (WCE). (B) HCT116 (WT), HCT116 hypomorphic DNMT1 (MT1 hypo), and HCT116 DNMT3B KO (3B KO) cells were treated with H_2O_2 at the indicated concentrations in mM for 30 minutes and total nuclear protein was collected. y-axis is SIRT1 over Actin levels relative to 8 mM treated WT cells. The data presented is the mean of three independent experiments +/- SEM. * $p < 0.05$ by t-test. (C) NCCIT cells were infected with non-specific shRNA (NS) or DNMT1 shRNA. After 72 hours, they were untreated (Unt) or treated with 1 mM H_2O_2 for 30 minutes (H_2O_2). Tight chromatin is the remaining protein in the chromatin pellet after extraction with 0.45 M NaCl buffer. Band densitometry values are displayed as the ratio of DNMT1 knockdown over NS knockdown for protein levels in H_2O_2 treated cells. The data presented is the mean of three independent experiments +/- SEM. * $p < 0.05$ by one-tail t-test. (D) NCCIT cells were transiently transfected with empty vector (EV) or c-Myc-tagged OGG1 (Myc-OGG1) plasmids for 48 hours followed by 1 mM H_2O_2 treatment and analysis as in (C). # Myc-OGG1. * $p < 0.05$ by one-tail t-test.

2.5.2 DNMT1 Functions Upstream of SIRT1 Recruitment to Chromatin Following Oxidative Damage

We next queried the interdependency of the tightening of SIRT1 and DNMT1 to chromatin after H_2O_2 treatment. SIRT1 knockdown or inhibition causes an increase in the fraction of DNMT1 tightly bound to chromatin after H_2O_2 treatment relative to non-specific shRNA or mock treated cells, respectively (Figure 2.2A-C). To examine these dynamics further, we utilized HCT116 cells genetically rendered hypomorphic for DNMT1 or fully deleted for DNMT3B (Rhee et al., 2002; Spada et al., 2007). While H_2O_2 treatment recruits SIRT1 to the nucleus in the WT and DNMT3B KO lines, there is a significant reduction in nuclear SIRT1 in the DNMT1 hypomorph cells, both by biochemical fractionation (Figure 2.1B) and immunofluorescence (Figure 2.2D). However, the residual increase in nuclear SIRT1 in DNMT1 hypomorph cells suggests that additional DNMT1-independent mechanisms may exist for changes in SIRT1 localization after H_2O_2 treatment. Furthermore, shRNA knockdown of DNMT1 in NCCIT cells significantly reduces the amount of SIRT1 that becomes tightly bound to chromatin after H_2O_2 treatment to 0.5 fold of control while simultaneously leading to an increase in γ-H2AX levels (Figure 2.1C). While these results suggest that tightening of DNMT1 and SIRT1 binding to chromatin after H_2O_2 treatment are dependent on each other, DNMT1 appears to be necessary for the increase in binding of SIRT1 to chromatin.

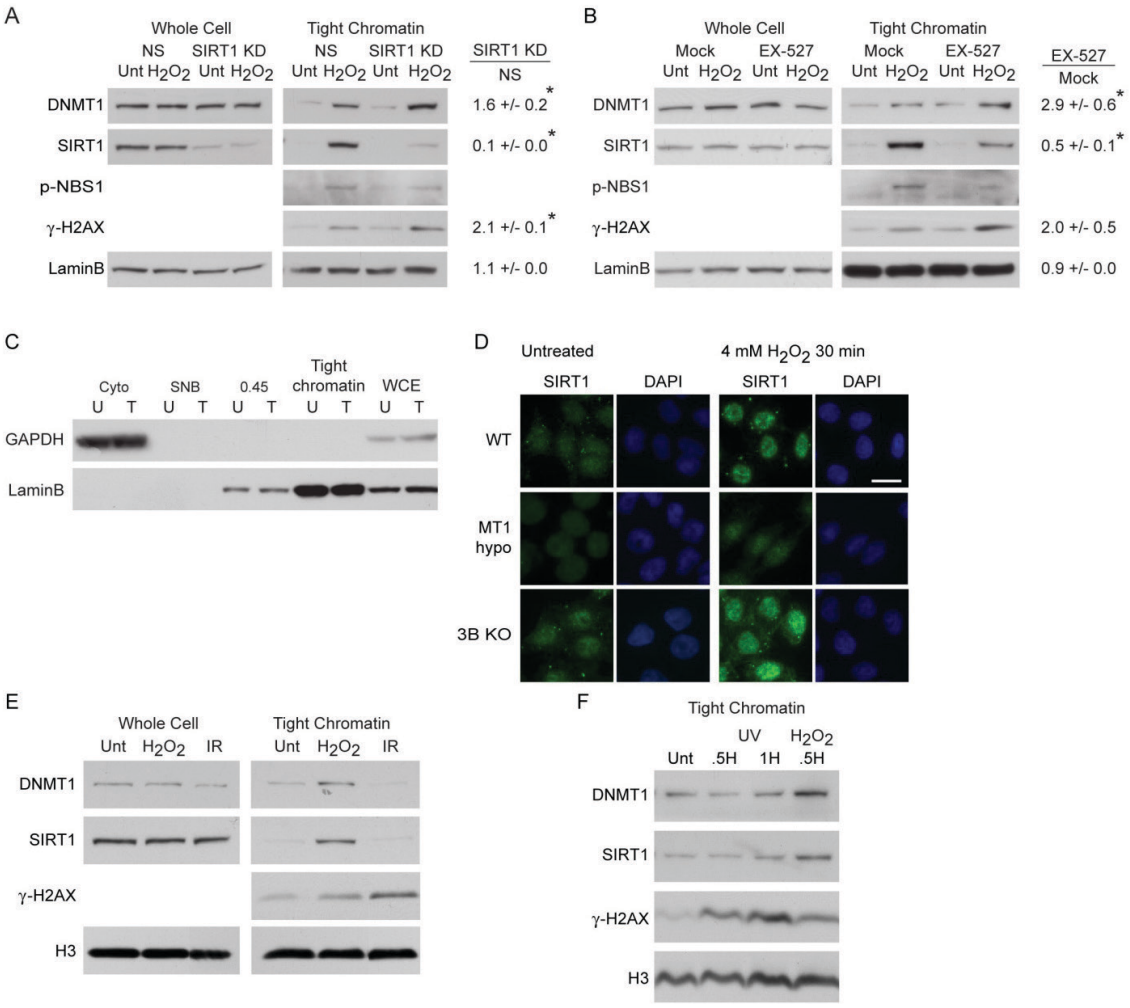


Figure 2.2. Modulation of tightening of DNMT1 and SIRT1 binding to chromatin. (A) NCCIT cells were infected with non-specific shRNA (NS) or SIRT1 shRNA. 72 hours after infection cells were treated with 1 mM H_2O_2 for 30 minutes. Tight chromatin is the remaining protein in the chromatin pellet after extraction with 0.45 M NaCl buffer. As a positive control for SIRT1 knockdown we demonstrate, following H_2O_2 treatment, reduced phosphorylated NBS1, which is known to be enhanced via SIRT1-mediated deacetylation (Yuan et al., 2007). Band densitometry is displayed as the value for H_2O_2 treated cells after SIRT1 knockdown divided by the value for H_2O_2 treated cells after NS knockdown. The data presented is the mean of three independent experiments +/- SEM. * indicates p-value < 0.05 by one-tail t-test. (B) NCCIT cells were pretreated with EX-527, which inhibits the deactylase activity of the enzyme for 32 hours followed by treatment with 1 mM H_2O_2 for 30 minutes (Solomon et al., 2006). Cell pellets were extracted and data analyzed as in (A). The data presented is the mean of three independent experiments

+/- SEM. * indicates p-value < 0.05 by a one-tail t-test. (C) Nuclear and cytoplasmic extracts from a representative tight chromatin extraction as in Figure 2.2, 2.1C, and 2.1D were blotted for a nuclear protein (LaminB) and a cytoplasmic protein (GAPDH). (D) WT, DNMT1 hypomorph (MT1 hypo), or 3B KO cells were treated with 4 mM H_2O_2 for 30 minutes. Cells were preextracted with detergent and then subjected to immunofluorescence analysis using antibodies against SIRT1. Nuclei are stained with DAPI. White scale bar is 5 mm. (E) NCCIT cells were untreated or treated with 1 mM H_2O_2 or 5 Gy IR and collected 30 minutes post-treatment. Cell pellets were extracted as in (A). Whole cell extracts and the tight chromatin fractions were analyzed by immunoblotting. (F) NCCIT cells were untreated or treated with 50 J/m^2 UV and collected 0.5 or 1 hour post-treatment, or treated with 1 mM H_2O_2 and collected 0.5 hours post treatment. Tight chromatin fractions were extracted as in (A).

Because ROS induces DNA damage in the form of base damage, single strand breaks, and double strand breaks, we next examined other types of DNA damaging agents and found that neither ionizing radiation nor ultraviolet light increase the tightness of binding of DNMT1 or SIRT1 to chromatin (Figure 2.2E&F). Additionally, inhibition of poly (ADP-ribose) polymerase (PARP), an enzyme involved in DNA repair of single and double strand breaks, or knockdown of key components of NER does not affect DNMT1 or SIRT1 recruitment to chromatin after H_2O_2 treatment (data not shown). Interestingly, we demonstrate that H_2O_2 treatment induces a significantly higher level of DNMT1 in the tight chromatin fraction in cells over-expressing c-Myc-tagged OGG1, the DNA glycosylase responsible for excising 8-oxo-dG during BER, compared to empty vector cells (Figure 2.1D). However, OGG1 overexpression does not affect the tightness of SIRT1 binding to chromatin possibly because the chromatin-bound levels of this protein are saturated after H_2O_2 treatment and therefore cannot be increased further by more DNMT1 recruitment.

2.5.3 DNMTs, SIRT1, and Polycomb Members Interact, as Part of a Large Multi-Protein Complex(es), after H2O2 Treatment

Having demonstrated that H_2O_2 treatment induces a DNMT1-influenced recruitment of SIRT1 to chromatin, we now queried whether such treatment might facilitate interactions between the two proteins and with other partners. We first observe, using DNMT1 co-immunoprecipitations from NCCIT cells after H_2O_2 exposure, a time-dependent interaction between endogenous DNMT1 and endogenous SIRT1 30 and 60 minutes after treatment (Figure 2.3A). We further validate this interaction by expressing a

FLAG-tagged full length DNMT1 in the DNMT1 hypomorph HCT116 cell line and finding that H_2O_2 treatment results in the interaction between the tagged DNMT1 protein and endogenous SIRT1 (Figure 2.3B).

We broadened our search for interacting proteins based on the previously mentioned association between DNA hypermethylated genes and PcG marks. In this regard, we observe that H_2O_2 treatment increases the interaction of DNMT1 with EZH2 and EED (Figure 2.3B). Further evidence for SIRT1 in this complex is demonstrated by immunoprecipitated endogenous nuclear SIRT1 pulling down increased levels of DNMT1, EZH2, SUZ12 and the PRC4 specific isoform of EED, EED2, after H_2O_2 treatment (Figure 2.3C & 2.4). Total cellular levels of the above proteins do not change after treatment (Figure 2.1A and data not shown). Initially, our entrée towards recognition of these interactions was based upon our previous demonstration that there is an acute recruitment of DNMT3B to an induced double strand break which is dependent on SIRT1 (O'Hagan et al., 2008). We now find that H_2O_2 treatment of NCCIT cells results in an endogenous interaction between SIRT1 and DNMT3B, as analyzed by immunoprecipitation of DNMT3B. This interaction is detectable within five minutes and increases up until 30 minutes after treatment (Figure 2.3D). As has been shown previously, DNMT3B and DNMT1 also interact (Kim et al., 2002) and this interaction does not change with treatment (Figure 2.3D). Our DNMT3B co-immunoprecipitation also reveals that the above interactions occur in the context of PRC4 because there is strong interaction of DNMT3B with EZH2, SUZ12 and EED2, both before and after H_2O_2 treatment (Figure 2.3D).

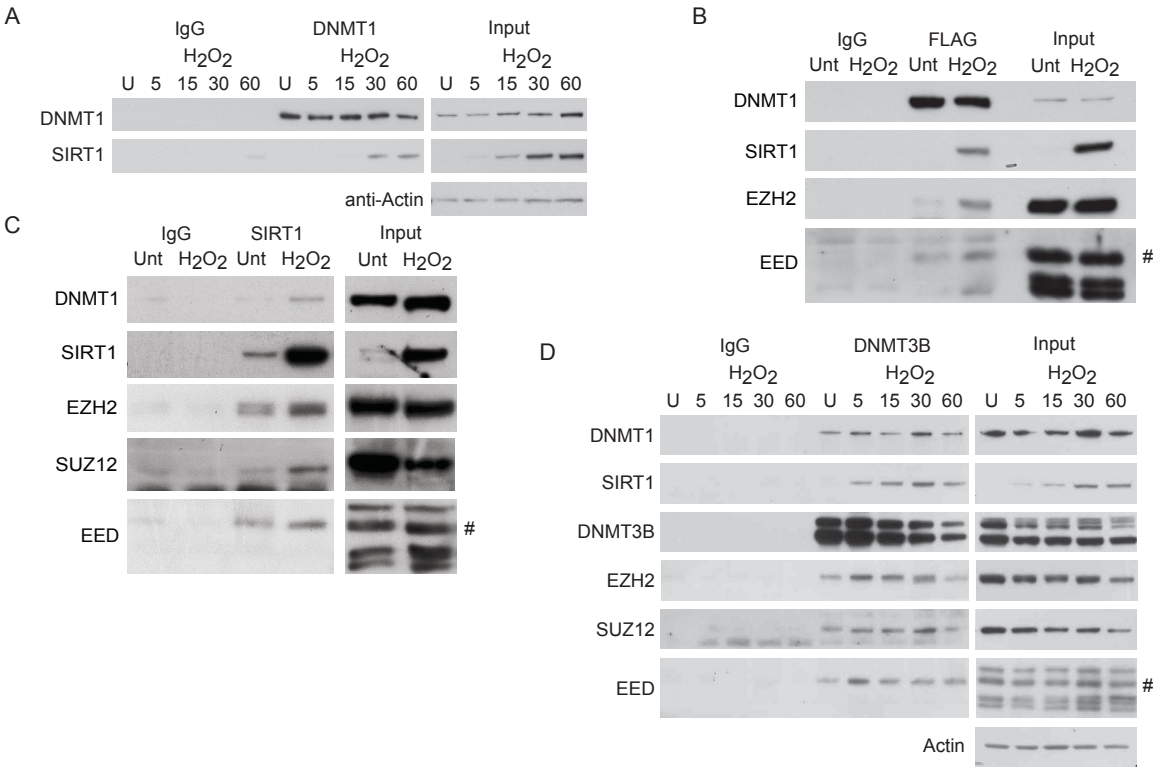


Figure 2.3. Oxidative damage induces the interaction between SIRT1, DNMTs, and PcG components.

(A) NCCIT cells were untreated or treated with 2 mM H_2O_2 and collected at the indicated time points in minutes after addition of H_2O_2 to the media. Co-immunoprecipitations were performed with control IgG or anti-DNMT1 antibodies. (B) HCT116 DNMT1 hypomorph cells expressing FLAG-DNMT1 were treated with 8 mM H_2O_2 for 30 minutes. Co-immunoprecipitations were performed using control IgG or anti-FLAG antibodies. # isoform 2 of EED. (C) NCCIT cells were treated with 2 mM H_2O_2 for 30 minutes and co-immunoprecipitations were performed using control IgG or anti-SIRT1 antibodies. # isoform 2 of EED. (D) NCCIT cells were treated as in (A) and co-immunoprecipitations were performed using control IgG or anti-DNMT3B antibodies. # isoform 2 of EED.

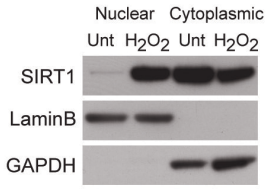


Figure 2.4. Validation of cytoplasmic and nuclear fractions for co-immunoprecipitations.

Nuclear and cytoplasmic extracts from the co-immunoprecipitation in Figure 2.3C were blotted for a nuclear protein (LaminB) and a cytoplasmic protein (GAPDH).

The triggering of multiple, individual, interactions suggests the H_2O_2 induced generation of a possible mega-complex(es). To examine this hypothesis, we separated nuclear protein complexes by size using sucrose gradient centrifugation from untreated and H_2O_2 treated cells. After H_2O_2 treatment, DNMT1, SIRT1 and DNMT3B and to a lesser extent SUZ12, EZH2, and EED proteins all migrate in the gradient in regions of higher molecular mass (greater than 650 kDa) than in untreated cells (Figure 2.5A). When we perform DNMT3B co-immunoprecipitation on pooled fractions from the gradient in untreated cells, the majority of the PcG proteins (EZH2, SUZ12, and EED2) and DNMT1 co-immunoprecipitate with DNMT3B in lower molecular weight pools 1 and 2 (smaller than 650 kDa – left panel - Figure 2.5B). Importantly, after treatment, the co-immunoprecipitation between DNMT3B and the PcG members and DNMT1 now shifts towards pools 2 and 3 indicating interaction of the proteins in a much larger complex(es) (middle panel - Figure 2.5B). After treatment, SIRT1 interacts with DNMT3B in the pools 2 and 3 (middle panel - Figure 2.5B), suggesting that all DNMT3B interacting members are present in the same size large complex(es) after H_2O_2 treatment.

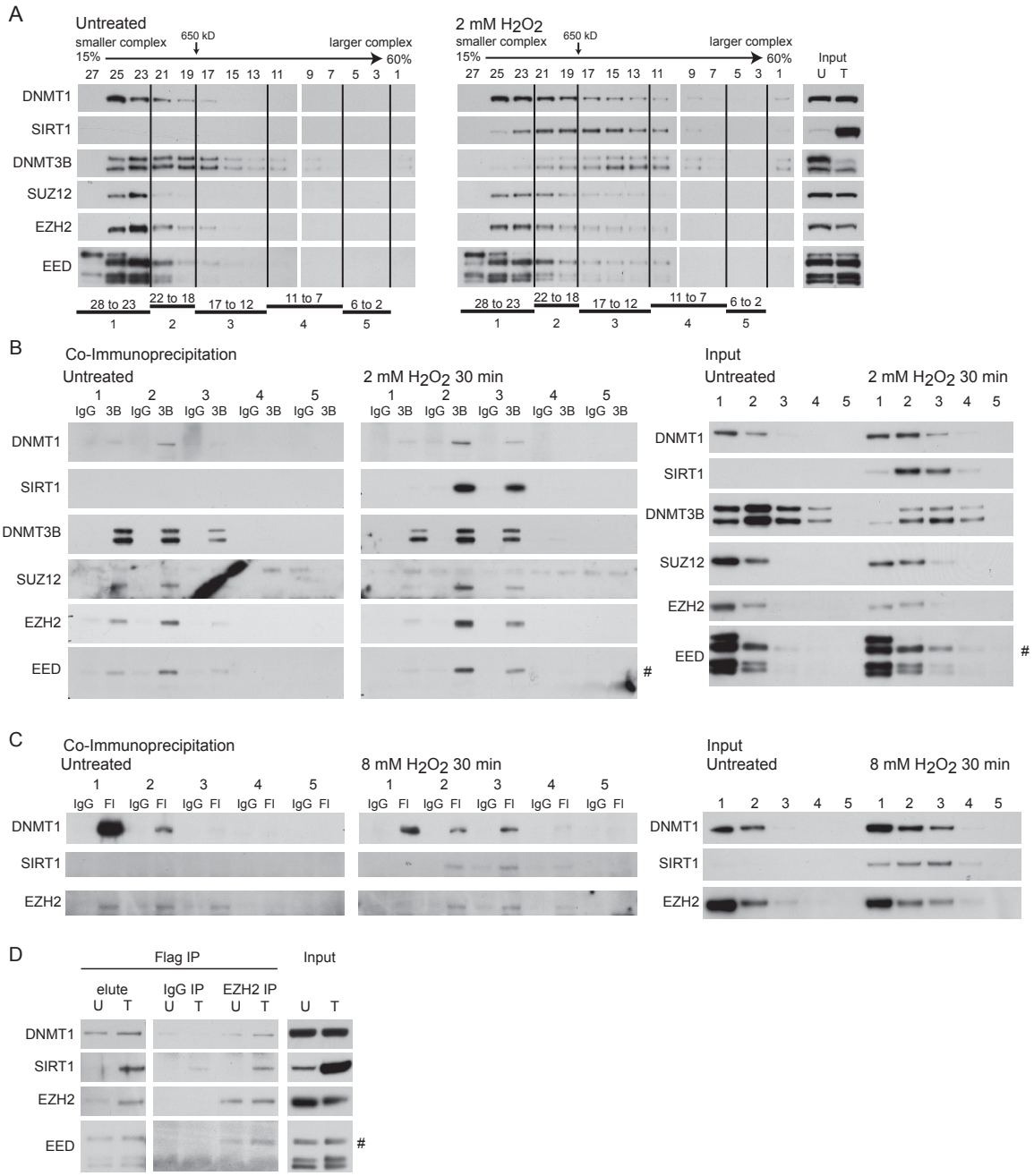


Figure 2.5. H_2O_2 treatment induces the formation of a large complex(es) containing DNA methyltransferases, SIRT1, and polycomb group proteins.

(A) Nuclear extracts from untreated NCCIT cells or cells treated with 2 mM H_2O_2 for 30 minutes were added to a 15 to 60% sucrose gradient and fractions were assayed by immunoblotting. Fraction numbers and 650 kDa molecular mass standard are across the top. Larger fraction numbers indicate smaller molecular weight of the complex(es). (B)

Fractions from (A) were pooled into 5 groups as indicated at the bottom of (A). Co-immunoprecipitations for control IgG or anti-DNMT3B (3B) antibodies were performed from the pooled fractions. Right panels are inputs from the pooled fractions. # isoform 2 of EED. (C) HCT116 DNMT1 hypomorph cells expressing FLAG-DNMT1 were untreated or treated with 8 mM H_2O_2 for 30 minutes. Nuclear extracts, sucrose gradients, and pooling of fractions were performed as in (A). Co-immunoprecipitations for control IgG or anti-FLAG (Fl) antibodies were performed from pooled fractions. Right panels are inputs from the pooled fractions. (D) Flag co-immunoprecipitations were performed in HCT116 DNMT1 hypomorph cells expressing FLAG-DNMT1 that were either untreated (U) or treated with 8 mM H_2O_2 for 30 minutes (T). After elution with flag peptide (elute) a second immunoprecipitation was done using IgG or EZH2 antibodies. # isoform 2 of EED.

To examine the above complex formation further, we performed similar gradient analyses and co-immunoprecipitations of DNMT1 in the DNMT1 hypomorph HCT116 cell line exogenously expressing FLAG-tagged full length DNMT1. The input protein analyses for this co-IP indicate that there is more nuclear DNMT1, SIRT1, and EZH2 in the higher molecular weight gradient fractions after H_2O_2 treatment than before treatment (right panel - Figure 2.5C). In untreated cells, the majority of FLAG-DNMT1 is immunoprecipitated from lower molecular weight gradient pools 1 and 2 and interacts with EZH2 in pools 1, 2, and 3 (left panel – Figure 2.5C). As expected, SIRT1 is not seen prior to treatment. After H_2O_2 treatment, FLAG-DNMT1 is now immunoprecipitated more prominently from pool 3 in addition to pools 1 and 2, peak interaction of EZH2 with FLAG-DNMT1 is now in pool 3, and interaction is seen with SIRT1 in pools 2 and 3 (middle panel – Figure 2.5C). Thus, as for DNMT3B, DNMT1 interacts with EZH2 and SIRT1 as part of a large multi-protein complex(es) after H_2O_2 treatment. To demonstrate that DNMT1, EZH2, and SIRT1 were indeed all bound together in these large complexes we performed sequential co-immunoprecipitation in the FLAG-DNMT1 cells. First, we immunoprecipitated DNMT1 containing complexes. After eluting these complexes, we immunoprecipitated EZH2 interacting proteins. As demonstrated by the SIRT1 band in the EZH2 IP lane in the DNMT1 elute from the treated cells, EZH2 that is bound to DNMT1 after H_2O_2 treatment is also bound to SIRT1 (Figure 2.5D). In total, this data suggests that oxidative damage induces the formation of a large complex(es) containing the DNMTs and PRC4 members.

2.5.4 DNMT1, SIRT1, and EZH2 Form DNA Damage Foci

To determine the possible interactions of DNMTs and members of PRC4 directly with DNA damage sites, we performed co-immunofluorescence of key interacting proteins with γ-H2AX in NCCIT cells. We used a paraformaldehyde fixation method that does not visualize DNA replication foci, since DNMT1 is constitutively present at these sites in S-phase (Figure 2.6A). After H_2O_2 treatment, the chromatin bound protein, LaminB, does not co-localize with γ-H2AX foci (Figure 2.6B). In contrast, there is an increase in total nuclear DNMT1, and in DNMT1 foci (0.3 foci in untreated and 4.1 in treated cells– Figure 2.7B), the majority of which co-localize with γ-H2AX (Figure 2.7C). The PcG member, EZH2, behaves similarly to DNMT1 (1.8 to 4.0 foci per cell) (Figure 2.7A-C). SIRT1, also, exhibits occasional larger foci after H_2O_2 treatment that co-localize with γ-H2AX (Figure 2.7A) but the considerable increase in nuclear staining makes precise quantitation difficult. Dense and widely distributed nuclear staining of DNMT3B before and after treatment did not allow visualization of foci after treatment. Overall, it appears that the tight binding of DNMT1 and SIRT1 to chromatin induced by H_2O_2, and the interaction between these two proteins and other PcG components, occurs, at least in part, at DNA damage sites.

2.5.5 Oxidative Damage Recruits Members of the H2O2 Induced Silencing Protein Complex to Promoter CpG Islands

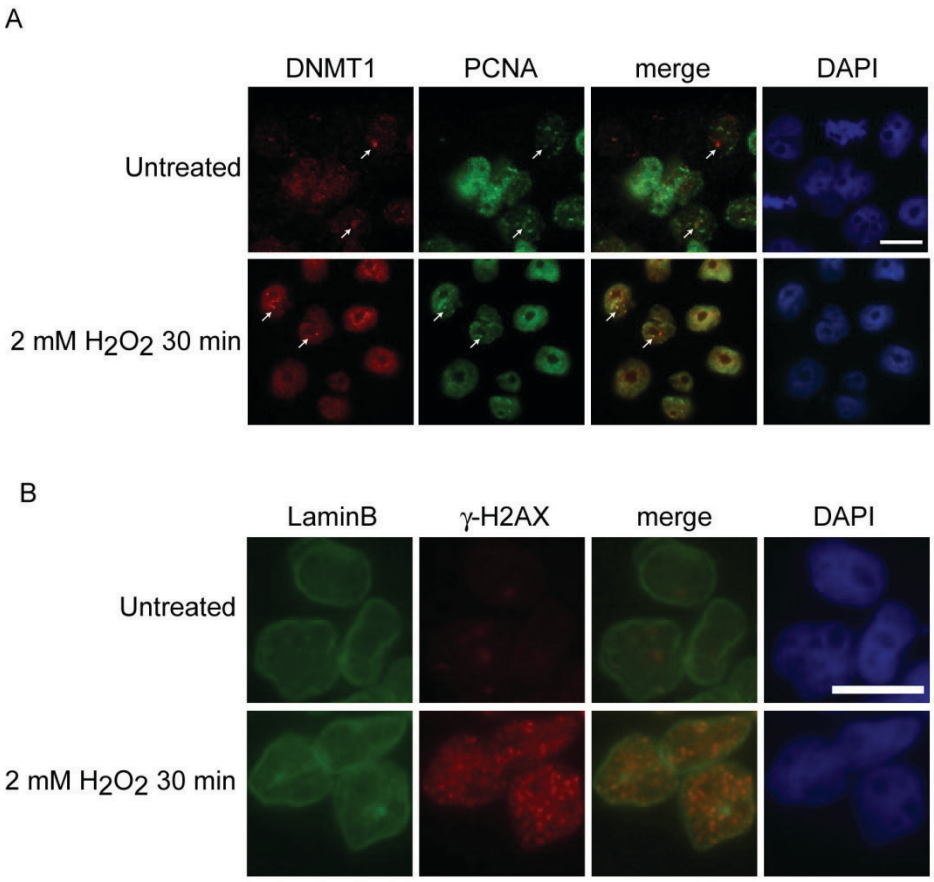


Figure 2.6. DNMT1 foci in late S phase cells.
(A) NCCIT cells were untreated or treated with 2 mM H_2O_2 for 30 minutes. Cells were preextracted with detergent, fixed with cold methanol and then subjected to immunofluorescence analysis using antibodies against PCNA and DNMT1. Nuclei were stained with DAPI. White arrows indicate examples of DNMT1 foci that co-stain with PCNA. (B) NCCIT cells were treated as in (A). Cells were preextracted with detergent and then subjected to immunofluorescence analysis using antibodies against LaminB and γ-H2AX. White scale bars are 5 mm.

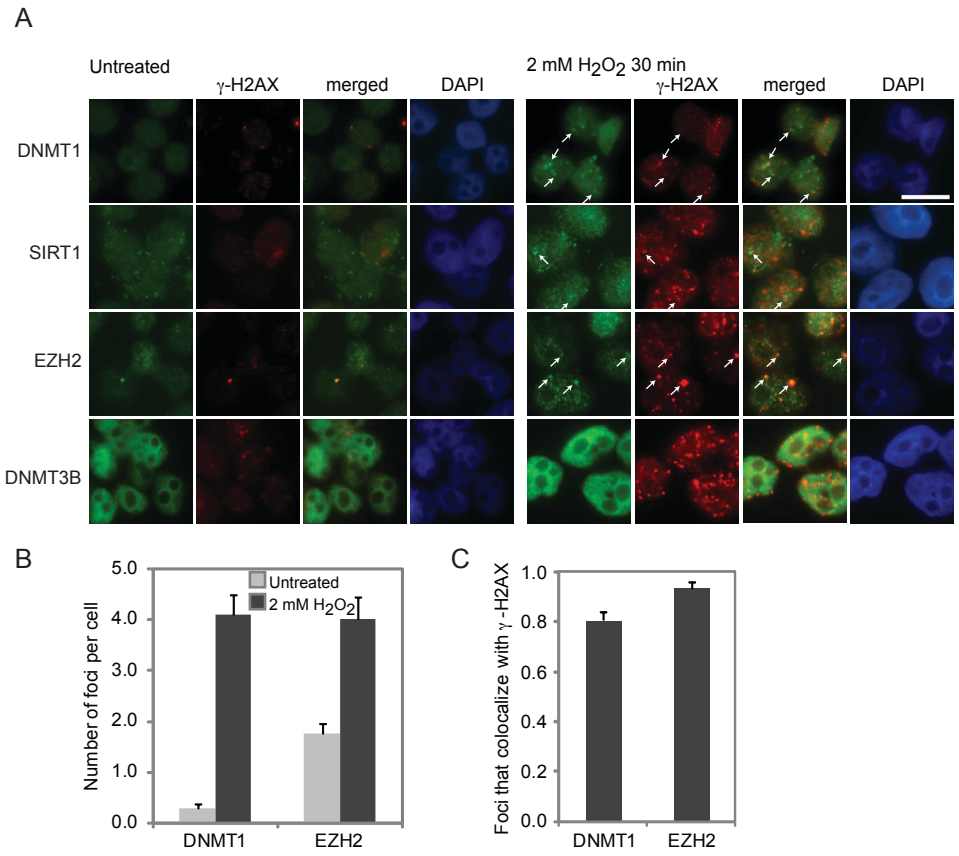


Figure 2.7. DNMT1 and EZH2 form nuclear foci after H_2O_2 treatment that co-localize with γ-H2AX.

(A) NCCIT cells were untreated or treated with 2 mM H_2O_2 for 30 minutes. Immunofluorescence analysis was performed using the indicated antibodies. White arrows indicate examples of foci. White scale bar is 5 mm. (B) More than 50 nuclei from cells in (A) were scored per antibody in at least two independent experiments. Graphs represent the sample mean +/- SEM. Grey and black bars are untreated and H_2O_2 treated cells, respectively. (C) More than 50 nuclei from cells in (A) were scored per antibody in at least two independent experiments. Graphs represent the sample mean +/- SEM. Black bars are H_2O_2 treated cells.

We next explored the genomic regions to which the oxidative damage-induced complexes may localize using the SW480 colon carcinoma cell line, where we have observed similar tightening of DNMT1 and SIRT1 to chromatin after H_2O_2 treatment (Figure 2.8A). We first find, utilizing customized ChIP-chip arrays for chromosome 18, 19, and 21 and histone H3 as a control, damage-induced concurrent shifts for DNMT1, DNMT3B, SIRT1 and γ-H2AX after H_2O_2 treatment (Figure 2.9A & 2.8B). In each case, the chromosome regions with enrichment constitute those previously mapped by others (Folle et al., 2010) to harbor high gene transcription activity and GC content. This enrichment shift is particularly observable for chromosome 21 where the sub-telomeric regions of the chromosome harbor most of the gene transcription and GC content (Figure 2.9A). Importantly, concomitant to the co-enrichments above, there is a notable loss of the same silencing proteins from the transcriptionally inactive and GC-poor regions (Figure 2.9A).

For all three chromosomes examined changes in ChIP signals are most prominent around the transcription start sites (TSS) of genes (example for DNMT1 signals in Figure 2.8C). To explore this finding in more depth, we used promoter arrays to examine the co-localization of the proteins at gene promoters across the genome. In these promoter studies we included, one of the key PcG members, EZH2. We also matched the results to previously obtained genome wide expression array data (Easwaran et al., 2010). These studies not only confirm targeting of all the tested proteins, after damage, to transcriptionally active GC-rich promoter regions, but also extend our findings in key ways.

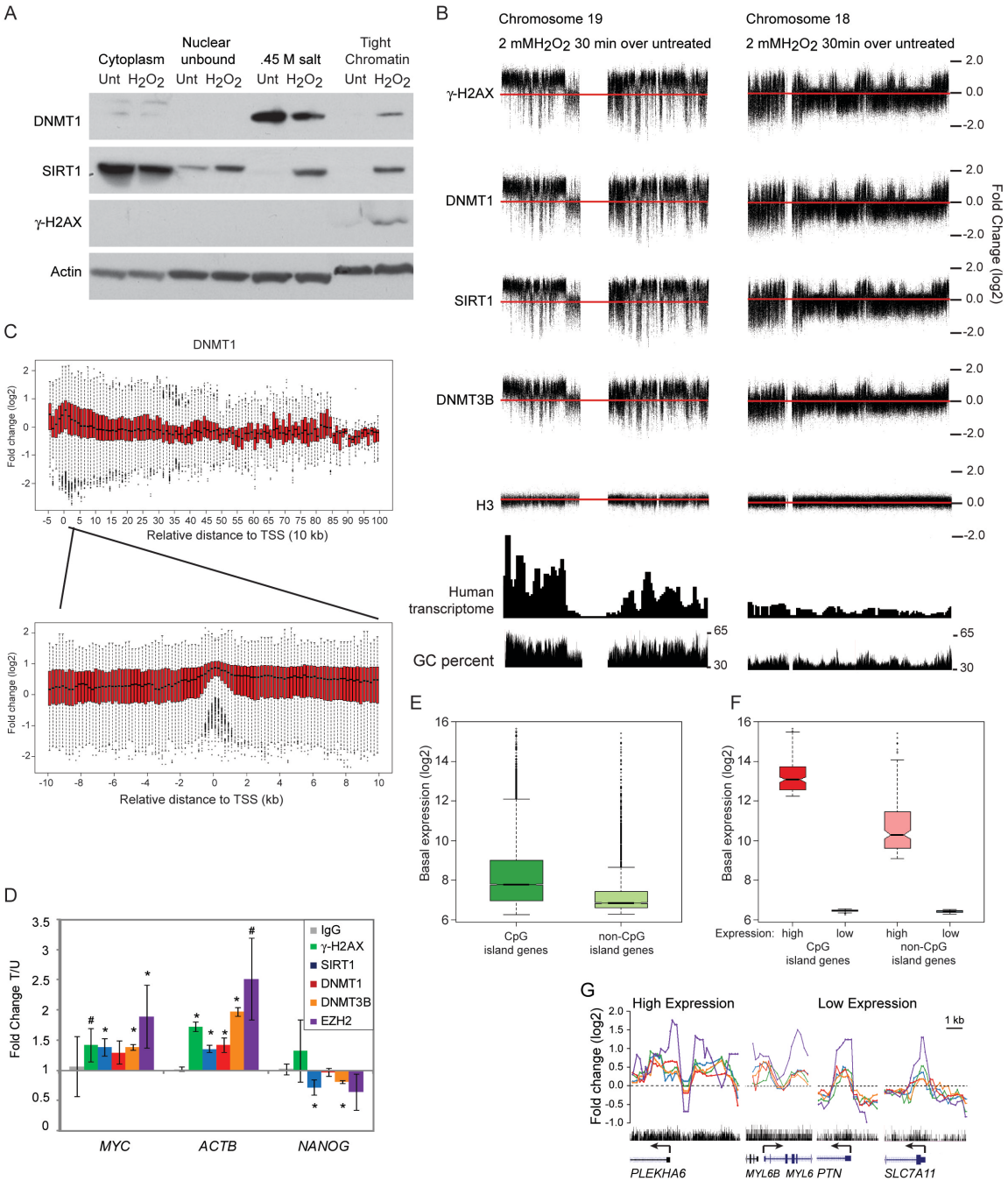


Figure 2.8. The silencing complex is targeted to regions of high transcription and GC content near TSSs.

(A) SW480 cells were either untreated or treated with 8 mM H_2O_2 for 30 minutes. Cell pellets were extracted sequentially with cytoplasmic buffer, soluble nuclear buffer, and 0.45 M NaCl buffer. Tight chromatin fraction is the remaining protein in the chromatin pellet after extraction with the 0.45 M NaCl buffer. Protein levels in each fraction were

analyzed by immunoblotting. Actin serves as a loading control. (B) SW480 cells were either untreated or treated with 2 mM H_2O_2 for 30 minutes, followed by ChIP for γ-H2AX, DNMT1, SIRT1, DNMT3B, or H3. Samples were hybridized to the 1M Agilent customized whole chromosome ChIP-chip array of chromosome 18, 19, and 21. Differences in log2 ratios of IP signals over input signals between treated and untreated samples were plotted along chromosome 19 and 18 for the respective antibodies. Red line represents the zero (no change) line. Signals above the red line represent gains of corresponding marker. Signals below the red line represent losses of corresponding marker. (C) Box plots were created by binning probes from chromosome 18, 19, and 21 on the 1M Agilent array described in (B) according to the distances to their corresponding TSSs. y-axis depicts the differences in log2 ratios of IP signals over input signals between the treated DNMT1 ChIP sample and the untreated sample. x-axis indicates the relative distances of each set of probes to their corresponding TSSs (units are in 10 kb for the upper panel and 1 kb for the lower panel). In the upper panel, the region from 50 kb upstream to 1000 kb downstream of the TSS is presented in 10 kb increments. In the lower panel, the region from 10 kb upstream to 10 kb downstream of the TSS is more tightly examined, with probes grouped in 200 bp increments. (D) SW480 cells were either untreated or treated with 2 mM H_2O_2 for 30 minutes, followed by ChIP for IgG, γ-H2AX, DNMT1, SIRT1, or DNMT3B. The y-axis is the ratio of the IP quantitative PCR values over input values between treated and untreated samples. The data presented is the mean of three independent experiments +/- SEM. * $p < 0.05$ and # $p < 0.1$ by one-sided t-test. (E) (F) Box plots for basal expression level of corresponding groups of genes analyzed in Figure 2.9C and 5E. y-axis depicts the log2 expression levels. In (E), dark green represents the group of CpG island genes, and light green represents the group of non-CpG island genes. In (F), red represents the group of CpG island genes with high expression, and pink represents the group of non-CpG island genes with high expression. Low expression genes for CpG island and non-CpG island genes are represented by box plots as indicated on the x-axis. (G) Plots of ChIP-chip signals for γ-H2AX (green), DNMT1 (red), SIRT1 (blue), DNMT3B (orange), and EZH2 (purple) in treated samples over untreated samples for individual genes. y-axis depicts the differences in log2 ratios of IP signals over input signals between treated and untreated samples. Black dash line represents the no change line. Black vertical bars indicate GC percent, ranging from 30% to 100%. Blue lines-boxes represent the position and construction of genes, with the boxes indicating the position of exons, lines indicating the position of introns, and arrow indicating the direction of transcription. The names of genes are indicated at the bottom of the plots.

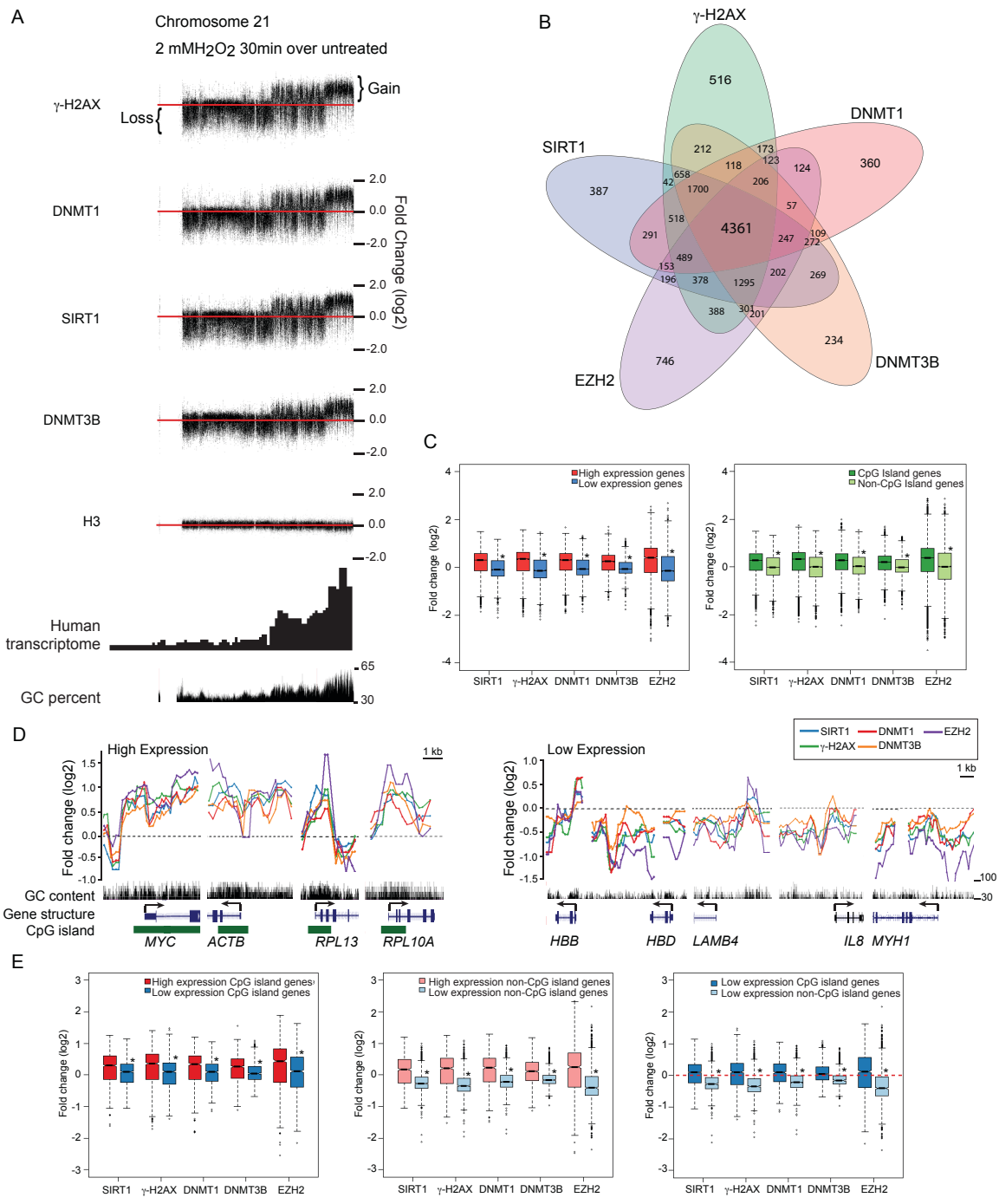


Figure 2.9. Oxidative damage induces recruitment of silencing proteins to the promoters of actively transcribed genes and/or high GC content regions.
(A) SW480 cells were either untreated or treated with 2 mM H_2O_2 for 30 minutes, followed by ChIP-Chip using the whole chromosome array of chromosomes 18, 19, and 21. Differences in log2 ratios of IP signals over input signals between treated and

untreated samples were plotted along chromosome 21. Red line represents the zero (no change) line. Signals above and below the red line represent gain and loss, respectively, of corresponding marker. (B) Venn diagram for ChIP enriched genes for each antibody in treated over untreated samples. SW480 cells were either untreated or treated with 2 mM H_2O_2 for 30 minutes. ChIP samples were hybridized to the 244K promoter array. (C) For box plots, the y-axis depicts the differences in log2 ratios of IP signals over input signals between treated and untreated samples. For the left panel, red and blue represent the top 1000 genes with high and low expression, respectively, from expression array data. In the right panel, dark and light green represent groups of CpG island and non-CpG island genes, respectively. * $p < 2\times10^{-10}$ by two-tail t-test. (D) Plots of ChIP-chip signals in treated over untreated samples for individual genes. y-axis is the same as in (C). Black dash line represents the no change line. Black vertical bars indicate GC content, ranging from 30% to 100%. Blue lines-boxes represent the position and construction of genes, with the boxes indicating the position of exons, lines indicating the position of introns, and arrow indicating the direction of transcription. Green boxes represent the position of CpG islands. The names of genes are indicated at the bottom of the plots. (E) Values were plotted as in (C). In the left panel, red and blue represent the top 100 CpG island genes with high and low expression, respectively. In the middle panel, pink and light blue represent the top 100 non-CpG island genes with high and low expression, respectively. In the right panel, blue and light blue represent the top 100 CpG island and non-CpG island genes, respectively, with similar levels of low expression. Red dash line represents the no change line.

First, DNMT1, DNMT3B, SIRT1, EZH2 and γ-H2AX enrichment that is lost from transcriptionally poor and low GC-content regions is translocalized to the promoters of genes that contain CpG islands and are highly expressed, with a high degree of overlap of enrichment between the different proteins (Figure 2.9B-E). The direct targeting to promoter CpG islands can be appreciated for example genes, *MYC*, *ACTB, RPL13,* and *RPL10A* (Figure 2.9D-left panel). The loss of enrichment for low expression, non-CpG island promoter genes is well appreciated for the example genes, *HBB, HBD, LAMB4, IL8,* and *MYH1* (Figure 2.9D-right panel). These genomic analyses were confirmed by local ChIP and quantitative PCR (Figure 2.8D).

Second, transcriptional activity is associated with targeting of the members of the complex separately from the presence of CpG islands (Figure 2.9E & 2.8E-F). We find that high expression genes gain more enrichment than low expression genes in groups of genes both with and without CpG islands (Figure 2.9E, left and middle panels). However, the presence of CpG islands is still important as targeting is increased in CpG island versus non-CpG island genes with similar low basal expression (Figure 2.9E, right panel & Figure 2.8F). Further scrutiny of the pattern of translocalization revealed that the position of peaks also correlates with areas of high GC content, including but not limited to CpG islands (Figure 2.9D [note gene *ACTB*] & Figure 2.8G), indicating that GC content, in addition to the presence of CpG islands, is a contributing factor to targeting. Altogether, this data suggests that members of the complex undergo H_2O_2-induced enrichment at gene promoters with high expression and/or high GC-content, including those with CpG islands.

2.5.6 Functional Consequences of Enrichment of Members of the Oxidative Damage-Induced Complex.

The fact that oxidative damage-induces proteins involved in gene silencing to form a complex and be enriched at CpG island-containing promoters, suggests a potential functional role of this complex in transcriptional changes accompanying DNA damage and/or cancer-specific, abnormal gene silencing. Examination of changes in histone marks, transcription, and DNA methylation supports this hypothesis.

We first examined histone modifications at CpG island-containing gene promoters most targeted by members of the complexes. Using genome-wide promoter arrays and histone H3 to normalize for nucleosome positions, we observed, as expected, marked relative enrichment for the active transcription marks, 3MeK4H3 and AcK16H4, and low amounts of the PcG repressive mark, 3MeK27H3, at the promoter of high expression genes in untreated cells (Figure 2.10A). After treatment, there is reduction in the active marks 3MeK4H3 and AcK16H4, the later is consistent with the deacetylation activity of SIRT1, and enrichment of the H3K27me3 mark, which can be catalyzed by the PcG component EZH2 (Figure 2.11A). These global changes are verified by examining patterns at the CpG island-containing promoters of genes, including, *MYC, ACTB, SFRP4, MLH1, SFRP5*, and *TIMP3* (Figure 2.11B) and by local ChIP studies (Figure 2.10B). We see similar changes, although with less magnitude, for high expression non-CpG island genes (Figure 2.10C left panel). Interestingly, there appears to be a slight gain in 3MeK4H3 and loss of 3MeK27H3 (Figure 2.10C, right panel) at low expression non-CpG island gene promoters where there is relative decrease in the complex constituents (Figure 2.9E, right panel).

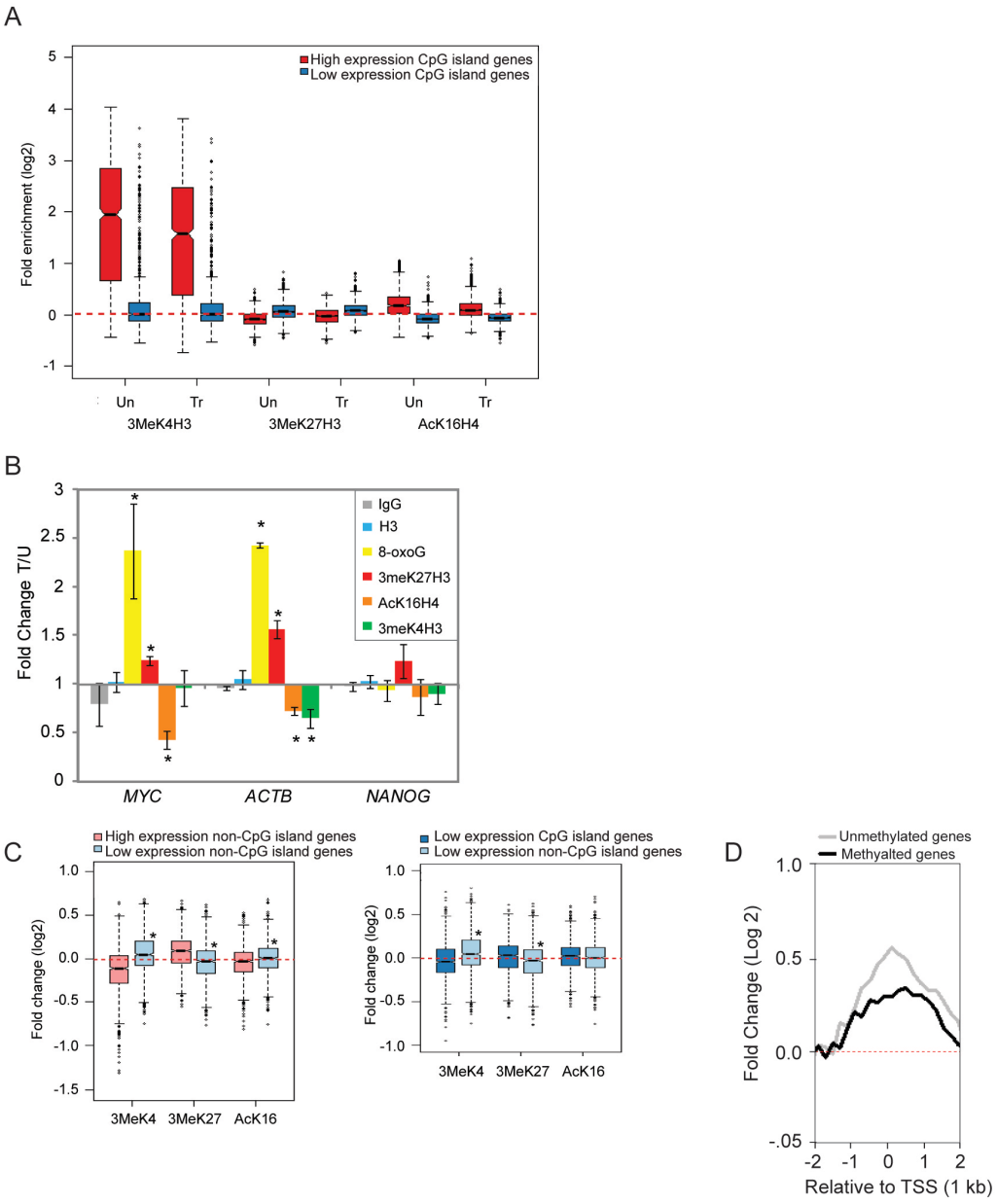


Figure 2.10. Histone mark changes after H_2O_2 treatment occur preferentially at promoters of highly expressed and CpG island containing genes.

(A) Box plots are constructed using the ChIP enrichment for probes within the range of 2000 bp upstream and 2000 bp downstream relative to corresponding TSSs. y-axis depicts the log2 ratios of IP signals over H3 signals. (B) SW480 cells were either untreated or treated with 2 mM H_2O_2 for 30 minutes, followed by ChIP for IgG, H3, 8-oxoguanine (8-oxoG), 3meK27H3, AcK16H4, or 3meK4H3. The y-axis is the ratio of the IP quantitative PCR values over H3 values between treated and untreated samples. The

data presented is the mean of three independent experiments +/- SEM. * $p < 0.05$ by one-sided t-test. (C) y-axis depicts the differences in log2 ratios of IP signals over H3 signals between treated samples and untreated samples. In the left panel, pink represents the top 100 non-CpG island genes with high expression, and light blue represents the top 100 non-CpG island genes with low expression. In the right panel, blue represents the top 100 CpG island genes with low expression, and light blue represents the top 100 non-CpG island genes with similar level of low expression. Red dash line represents the no change line. * indicates a p-value of less than 2×10^{-10} for the corresponding antibody when performing a two-tail t-test for the two groups. (D) SW480 cells were either untreated or treated with 2 mM H_2O_2 for 30 minutes, followed by ChIP for DNMT1. Samples were hybridized to the Agilent promoter 244K ChIP-chip array. ChIP-chip signals were compared between 293 unmethylated genes in SW480 (grey) and 269 methylated genes in SW480 (black) that commonly undergo cancer specific DNA methylation. Plots were generated by binning probes for the two groups of genes according to the distances to their corresponding TSSs in 200 bp increments. The median value for each bin is then graphed for each group of genes. y-axis depicts the differences in log2 ratios of IP signals over input signals between the treated and the untreated sample. x-axis indicates the relative distances of each set of probes to their corresponding TSSs. The difference between the two plots is statistically significant between -800 kb to +2000 kb of the TSS (except for 1000 to 1200 kb) using two-sided Wilcoxin rank sum tests with p-values of < 0.05.

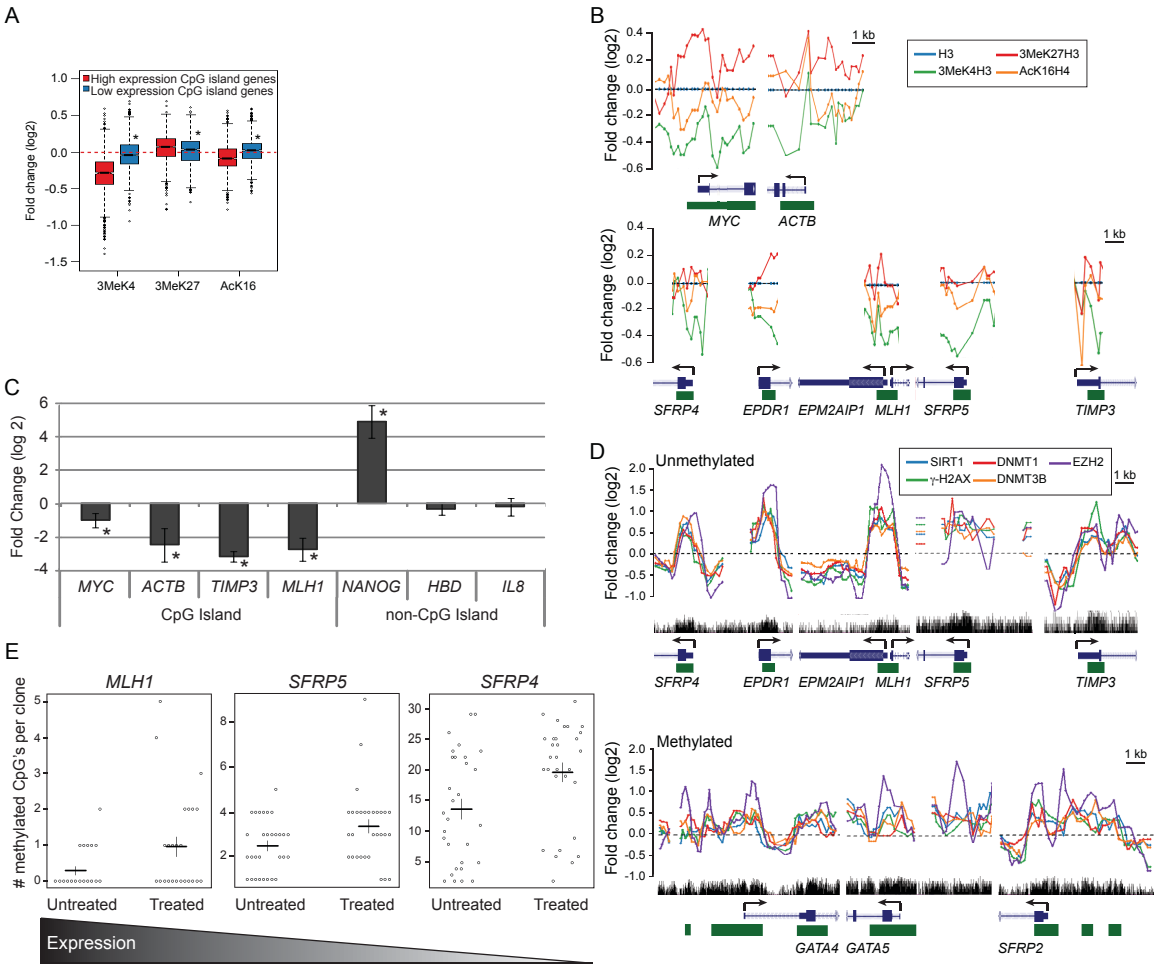


Figure 2.11. Gene promoters with oxidative damage-induced enrichment of the members of the silencing complex have reduced levels of nascent transcription and/or increased CpG methylation.

(A) ChIP samples were hybridized to the 1M promoter array. Plots are of ChIP-chip signals from SW480 cells treated with 2 mM H_2O_2 for 30 minutes over untreated samples. Box plots are constructed as in Figure 2.9C except ChIP over input signals are normalized to H3. Red and blue represent the top 100 CpG island genes with high and low expression, respectively. * $p < 2\times10^{-10}$ by two-tail t-test. (B) Plots are constructed as in Figure 2.9D with ChIP signals normalized to H3. (C) SW480 cells were either untreated or treated with 2 mM H_2O_2 for 30 minutes and the nascent RNA was labeled concurrently. Quantitative RTPCR data are presented as the mean of the log 2 ratio of the treated over the untreated values for three independent biological replicates +/- SEM. * $p < 0.05$ by one-sided t-test. (D) Plots are constructed as in Figure 2.9D. (E) Bisulfite sequencing was performed on DNA from SW480 cells that were untreated or treated with 2 mM H_2O_2 for 30 minutes. Results are shown as a combination of three independent

biological replicates with at least eight clones per experiment. Circles are the individual clones. Black horizontal lines are the mean for all clones with the vertical line representing the standard error. p-values by one-sided Welch's t-test for MLH1, SFRP5, and SFRP4 are 0.014, 0.012, and 0.005, respectively. The gradient bar at the bottom depicts the relative log 2 basal expression levels, by expression array, which are 11.13, 6.82, and 6.40 for *MLH1*, *SFRP5*, and *SFRP4*, respectively.

Second, the change in 3MeK4H3 levels suggests that enrichment of members of the complex may also rapidly induce repressive transcriptional changes mainly in CpG island-containing genes with significant basal expression. Indeed, the CpG island-containing genes, *MYC*, *ACTB*, *TIMP3*, and *MLH1*, all have reduced nascent transcription levels within 30 minutes after treatment (Figure 2.11C). In contrast, the non-CpG island-containing genes, *NANOG*, *HBD*, and *IL8*, either gain or have no change in their nascent transcription levels suggesting that the reduction in transcription for high expression genes is not solely due to a global DNA damage-induced decrease in transcription (Figure 2.11C).

Third, we find important correlates for genes that develop cancer-specific, abnormal, CpG island DNA hypermethylation. Genes with a high frequency for this change in colon cancer that are unmethylated or hypermethylated in SW480 cells, have damage-induced enrichment of the silencing proteins (Figure 2.11D & Figure 2.10D).

Fourth, we examined DNA methylation by bisulfite sequencing for representative genes. We did not observe any CpG methylation in the promoter CpG island of the high expression gene, *MYC,* which does not become hypermethylated in cancer (data not shown). However, at the short time point examined we find increases in DNA methylation for the *MLH1*, *SFRP5*, and *SFRP4* genes, which are frequently hypermethylated in colon cancer but are unmethylated in SW480 cells (Figure 2.11E). These three genes are basally expressed to varying degrees in SW480 cells and the level of observed increases in DNA methylation are inverse to this expression (Figure 2.11E). This finding fits well with emerging data that cancer specific, promoter DNA methylation mostly targets genes, which normally have low basal expression (Hahn et al., 2008). This

data suggests that localization of members of the complex to gene promoters has functional consequences, including histone mark changes, reduction in nascent transcription, and/or increases in DNA methylation.

2.5.7 Inflammation Induced Changes in EZH2, SIRT1, and DNMT1 in a Mouse Model of Colitis.

To examine whether our *in vitro* findings are applicable to the cancer risk state of inflammation, we studied an *in vivo* model of colitis in which infection with the human commensal enterotoxigenic *Bacteroides fragilis* (ETBF) induces inflammation and tumorigenesis primarily in the distal colons of Multiple intestinal neoplasia (*Min*) mice which are heterozygous for loss of *Adenomatous polyposis coli* (*Apc*) (Rhee et al., 2009; Wu et al., 2009). Importantly, *in vitro* treatment of colon cancer cells with purified *B. fragilis* toxin induces an increase in γ-H2AX and ROS suggesting that this model provides an *in vivo* scenario to assess the endogenous impact of the oxidative damage we examined in our *in vitro* model (Goodwin et al., 2011). First, we examined whether any of the players involved in our complex become more tightly bound to chromatin in the inflamed distal tissue (Figure 2.12A). Similar levels of Villin, an epithelial marker, establish that cells harvested by scraping from ETBF and mock inoculated (sham) mice have similar epithelial content, though this method does not obtain pure populations of epithelial cells (Figure 2.12B). Using the salt gradient extraction employed in Figure 2.1A, while we detect no change in DNMT1 (data not shown), we demonstrate that both SIRT1 and EZH2 are more tightly bound to chromatin in distal, but not proximal, colon epithelial cells from ETBF mice than sham mice two days after inoculation (Figure 2.13A,

2.12B & 2.12C). These results suggest the change in binding is due to the specific high level of inflammation that occurs during the acute phase of the ETBF model.

Secondly, we examined whether any of the proteins in our complex interact in the inflamed tissue. We performed co-immunoprecipitations using anti-EZH2 antibodies in proximal or distal colon epithelial cells from two separate pairs of sham and ETBF mice. While in all cases EZH2 co-immunoprecipitates predominantly isoform 1 of EED (isoform is determined by comparison to the four isoforms present in mouse embryonic carcinoma cells - Figure 2.12D), co-immunoprecipitation of DNMT1 by EZH2 is more prominent in the tissue from the distal colon of the ETBF mice (Figure 2.13B & 2.12D) suggesting that this interaction is increased by the inflammation in this tissue.

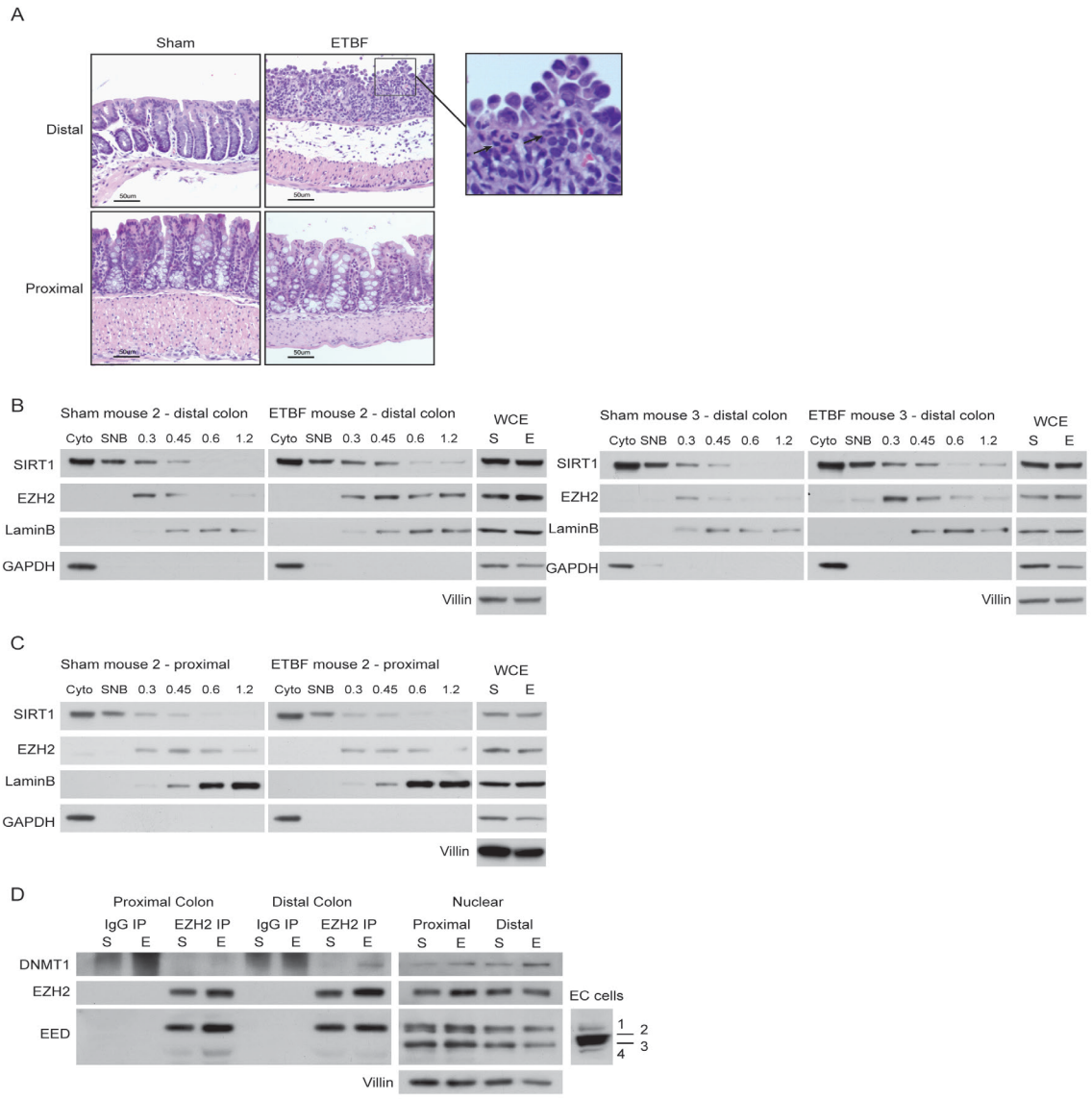


Figure 2.12. Mice infected with ETBF have inflammation in their distal colon. (A) H&E-stained tissue sections of proximal and distal colon from sham and ETBF-colonized mice 2 days after inoculation. Histopathology of the distal colon from the ETBF-colonized mouse demonstrates mucosal thickening, sub-mucosal edema, and inflammatory cell infiltration with rounding and detachment of enterocytes. Arrows indicate infiltrating neutrophils. Black scale bars are 50 mm. (B) Salt gradient extractions of distal colon epithelial cells from two mice each of Sham and ETBF used in the quantification in Figure 2.13A. Mice were mock colonized (Sham) or colonized with ETBF. Two days post-inoculation colon epithelial cells were collected by scraping, washed with PBS, and extracted sequentially using cytoplasmic extraction buffer (Cyto), soluble nuclear buffer (Sol), and buffers with increasing NaCl concentration. Whole cell

lysate was prepared separately (WCE). Protein levels in each fraction were analyzed by immunoblotting. (C) Salt gradient extractions of proximal colon epithelial cells from the second set of Sham and ETBF mice used in (B). (D) Biological replicate of EZH2 co-immunoprecipitation from Figure 2.13B from a second set of mice. Epithelial cells were collected as in (B). Co-immunoprecipitations for control IgG or anti-EZH2 antibodies were performed from nuclear extracts where chromatin bound proteins had been released by polyamine treatment. Cell extract from mouse embryonic carcinoma cells (EC cells) contains all four isoforms of EED and can be used to determine the isoforms present in the colon epithelial cells.

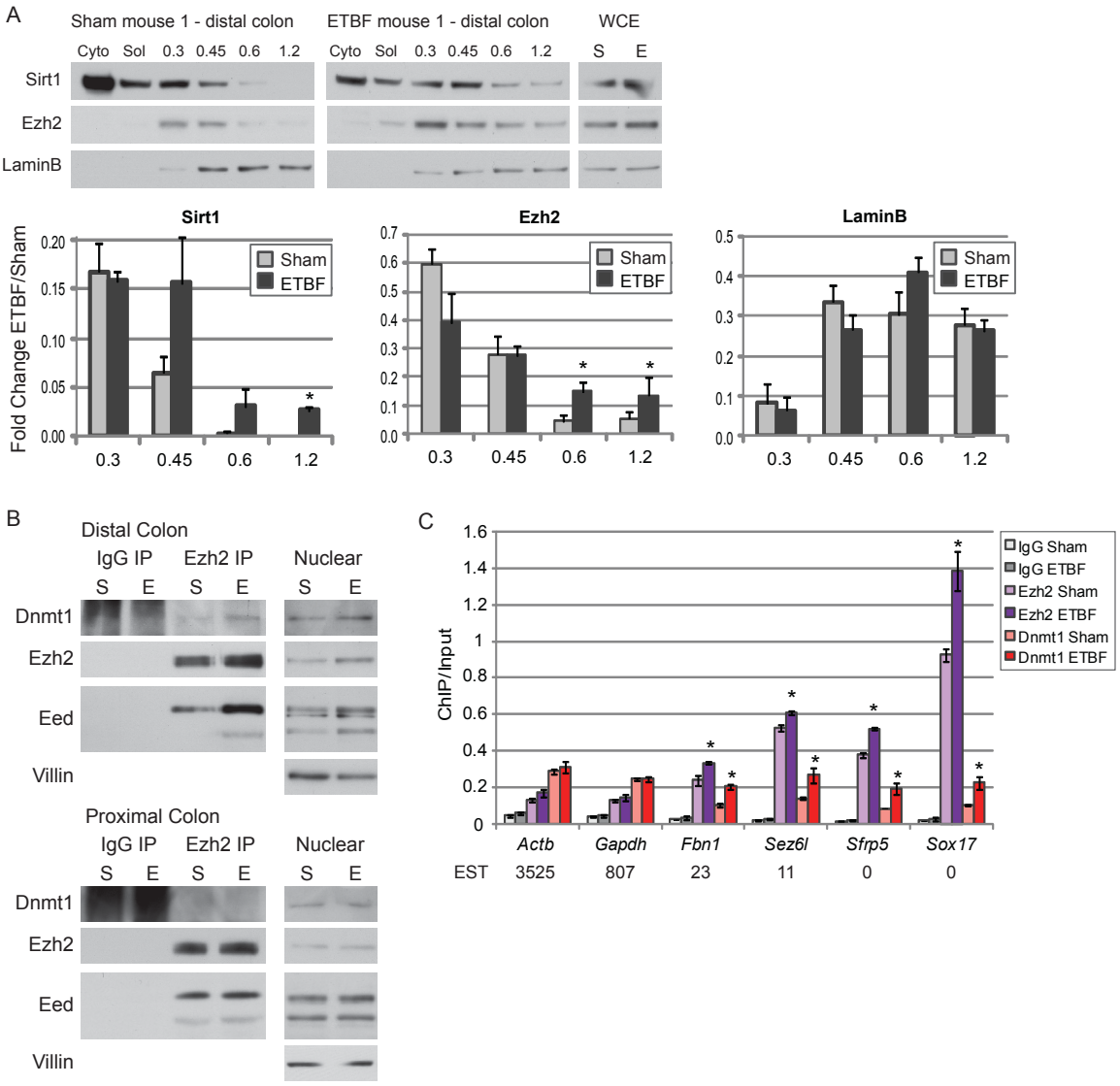


Figure 2.13. In a mouse model of acute colonic inflammation, members of the silencing complex become enriched at promoter CpG islands of low expression genes.

(A) Mice were sham-inoculated (sham) or inoculated with ETBF. Two days post-inoculation, colon epithelial cells were extracted sequentially using cytoplasmic extraction buffer (Cyto), soluble nuclear buffer (Sol), and buffers with increasing NaCl concentration. Whole cell lysate was prepared separately (WCE). Blots from one set of representative mice are depicted. The value calculated for each fraction is the ratio of that fraction over the total of all fractions. The graphs represent the mean values for 3 separate mice +/- SEM. * p-value < 0.05 by one-tail t-test. (B) Co-immunoprecipitations for control IgG or anti-EZH2 antibodies were performed in nuclear lysates of colon epithelial

cells from sham-inoculated mice (S) or ETBF-inoculated mice (E). Blots from one set of representative mice are depicted. (C) Using distal colon epithelial cells, ChIP was performed for IgG, EZH2, or DNMT1 and analyzed by quantitative PCR. The data presented is the mean of ChIP performed in samples from 3 sham and 3 ETBF mice +/- SEM. * $p < 0.05$ by one-sided t-test for the difference between the means. Values below the gene names are the expressed sequence tag (EST) counts for mouse intestine from the Unigene database.

Thirdly, we performed local ChIP for EZH2 and DNMT1 in epithelial cells from the distal colon of ETBF and sham mice. Interestingly, unlike the *in vitro* model, we do not see enrichment of EZH2 and DNMT1 at all promoter CpG islands (Figure 2.13C). High expression housekeeping genes such as *Actb* and *Gapdh* have no change in EZH2 or DNMT1 enrichment between the ETBF and sham mice. However, lower expression genes, such as *Fbn1*, *Sez6l, Sfrp5*, and *Sox17*, that have higher basal levels of EZH2 than the high expression genes, all have more enrichment of EZH2 and DNMT1 at their promoter CpG islands in inflamed distal epithelial cells from ETBF mice compared to sham mice. Interestingly, three of these genes, *Fbn1*, *Sez6l,* and *Sox17*, have been demonstrated to undergo inflammation and tumor-specific DNA methylation in a model of intestinal inflammation and all four are methylated in human cancers (Hahn et al., 2008). These data suggest that, in this model, where changes in ROS are likely less dramatic and induced over a longer time frame than for our *in vitro* model, the recruitment of members of the silencing complex is most persistent at promoter CpG island-containing genes with lower basal expression. These genes are the most likely to be targets of cancer-specific DNA methylation (Hahn et al., 2008).

2.6 Discussion

In the present study, we link several proteins involved in transcriptional repression to the DNA damage response. We provide evidence for a role for DNMT1 in the response of cells to H_2O_2 treatment. This enzyme becomes more tightly bound to chromatin after H_2O_2 treatment in the context of damage-induced foci that co-localize with γ-H2AX. Moreover, DNMT1 appears responsible for the tightening of the PRC4 component, SIRT1, to chromatin after H_2O_2 treatment. SIRT1 has been implicated in the response to DNA damage and transcriptional repression in many ways as previously discussed. While in inflamed mouse tissue we do not see an increase in binding of DNMT1 to chromatin, we see an increase in binding of the PRC4 components, SIRT1 and EZH2. Increases in binding of DNMT1 to chromatin may occur at an earlier time point than the one we studied, as the *in vivo* model time point is two days after infection compared to 30 minutes after treatment for our *in vitro* model. However, we do implicate DNMT1 in the *in vivo* response to colonic inflammation by demonstrating interaction between EZH2 and DNMT1 and enrichment of both of these proteins at the CpG islands of low expression genes in this inflamed tissue, in a manner consistent with our *in vitro* results.

Our observations presented here suggest that tightening of DNMT1 and SIRT1 to chromatin after H_2O_2 treatment is actively associated with DNA damage and/or repair. Our evidence by immunofluorescence studies for co-localization of DNMT1, SIRT1, and EZH2 to DNA damage-induced foci marked by γ-H2AX suggests targeting to ongoing DNA damage. Previously, we have implicated DNMT1 in double strand break repair and recently another group has further studied this role for DNMT1 (Ha et al., 2010; O'Hagan

et al., 2008). We did not see DNMT1 become more tightly bound to chromatin after IR treatment and the observed narrow localization of γ-H2AX signal at the promoter region by ChIP-chip after H_2O_2 exposure is not representative of the typical mega base domain enrichment seen after double strand breaks (Rogakou et al., 1999). Thus, our findings suggest that, after H_2O_2 treatment, double strand breaks are not the dominant trigger for the dynamics we are reporting. Because levels of OGG1 modulate the tightening of DNMT1 binding to chromatin, we suggest that either the specific 8-oxodG base damage or the BER pathway that repairs this type of damage is responsible for recruitment of member of the silencing complex to chromatin. We propose that the promoter targeted γ-H2AX is likely marking sites of base damage and this hypothesis is supported by the enrichment of 8-oxoguanine at GC-rich promoters after H_2O_2 treatment. We suggest that the enrichment of damage at these areas of high GC content may be because guanine is the most easily oxidized of the four deoxyribonucleosides and therefore may be targeted by oxidative damage (Steenken, 1997). The co-enrichment of DNMT1, SIRT1, DNMT3B, and EZH2 at these sites suggests that these proteins are being localized to sites of base damage induced by H_2O_2 treatment. While in our *in vivo* model we have not assessed whether the enrichment of key complex constituents is also occurring at the sites of base damage, it has been demonstrated that purified *B. fragilis* toxin causes an increase in γ-H2AX and ROS in colonic epithelial cells. This data suggests that our *in vivo* model involves a similar induction of oxidative damage as our *in vitro* model (Goodwin et al., 2011).

One of the intriguing implications of our data is the potential role for increased levels of cellular ROS that accompany cancer-risk states such as inflammation in the formation

of cancer-specific aberrant patterns of DNA methylation and transcriptional silencing. First, as we have noted, cellular transformation has been associated with the presence of the PRC4 iso-complex (Kuzmichev et al., 2005) that we now link to DNMTs during H_2O_2 exposure. This complex has altered substrate specificity from the typical PRC2/3 complexes possibly due to the specific isoform of EED that it contains (EED2) (Kim et al., 2007). Since both pre-neoplastic and transformed cells undergo a significant amount of endogenous oxidative damage (Federico et al., 2007), the basal PRC4 complex previously described in transformed cells may be the same as the complex we describe here. Because nuclear SIRT1 levels increase after H_2O_2 treatment, it is possible that the increase in interaction demonstrated between SIRT1 and the other proteins in Figure 2.3 is due to higher levels of SIRT1, not to an induced interaction per say. However, either cause has the same outcome, namely higher levels of the complex.

Second, our findings suggest one potential mechanism that might help explain a conundrum in the abnormalities of DNA methylation in human cancer - namely, why cancer cells simultaneously harbor both widespread chromosomal loss of DNA methylation and increased DNA methylation in CpG islands of gene promoters (Jones and Baylin, 2007). In terms of the losses, we find that enzymes that catalyze DNA methylation, DNMT1 and DNMT3B, shift away from non-GC-rich gene and chromosome regions. In a similar manner, it has previously been demonstrated in yeast and mammalian cells that DNA damage leads to a shift in localization of SIRT1 from repressed gene regions to sites of induced DNA damage resulting in transcriptional de-repression of genes that are basally repressed by SIRT1 (Mills et al., 1999; Oberdoerffer et al., 2008). We suggest that when cells are exposed to chronic oxidative damage that is

present during all phases of tumorigenesis, the induced shifts in chromosome localization that we demonstrate may be associated with losses of DNA methylation observed in cancer cells.

Finally, our observations may also help explain gains in DNA methylation at gene promoters in cancer cells. By examining histone mark and transcription changes, we demonstrate that enrichment of members of this silencing complex is associated with gene silencing. Importantly, however, in cells progressing towards transformation, the above translocalization would probably be transient at most genes with high basal transcription levels, such as housekeeping genes, and oncogenes, for which silencing would be detrimental to tumor cell growth. This hypothesis is supported by the lack of EZH2 and DNMT1 enrichment at the promoters of high expression CpG island-containing genes in the less harsh, longer time-frame, inflammatory *in vivo* model studied. In this regard, active transcription may prevent *de novo* promoter CpG island methylation (Thomson et al., 2010). We demonstrate that genes with a history of frequent, cancer specific, CpG island promoter DNA hypermethylation show damage-induced enrichment for the members of the complex in cell culture and enrichment of EZH2 and DNMT1 in inflamed mouse tissue. *In vitro* we see an increase in DNA methylation that correlates with the low basal expression level of these genes, which harbor PcG complexes in embryonic stem and progenitor cells (Ohm et al., 2007; Schlesinger et al., 2007; Widschwendter et al., 2007). We hypothesize that such localization of the DNMT-PRC4 complex and increase in DNA methylation at low expression promoter CpG island-containing genes might be more persistent over the course of chronic ROS damage during tumorigenesis setting up a scenario for the expansion of DNA methylation in the

CpG islands involved. Our previous work with a promoter CpG island, double strand break DNA damage model, suggests this time-dependent context for the expansion of such DNA hypermethylation (O'Hagan et al., 2008).

2.7 Experimental Procedures

2.7.1 Antibodies

- For western blot: DNMT1 (Sigma), SIRT1 (Delta Biolabs, Millipore), γ-H2AX (Millipore), EZH2 and Phopho-NBS1 (Cell Signaling), and SUZ12 (Abcam), and DNMT3B, Actin, and EED (Sigma).
- For immunofluorescence: DNMT1, PCNA, and LAMB (Santa Cruz Biotechnology), SIRT1 (Millipore), EZH2 (Cell Signaling), γ-H2AX (Millipore). Secondary antibodies were FITC anti-Rabbit (Invitrogen) and Alexa Fluor 555 anti-Mouse (Invitrogen).
- For co-immunoprecipitation: DNMT3B (described previously (Rhee et al., 2002)), SIRT1 (Millipore), FLAG (Sigma), Ezh2 (Cell Signaling), Rabbit and Mouse IgG (Millipore).
- For ChIP: DNMT3B (described previously (Rhee et al., 2002)), DNMT1 (Sigma), SIRT1 (Millipore), EZH2 (Cell Signaling), γ-H2AX (Millipore), H3 (Abcam), 3MeK4H3 (Millipore), 3MeK27H3 (Millipore), AcK16H4 (Millipore), and 8-oxoguanine (Millipore).
- For ChIP in mouse tissue: EZH2 (Cell Signaling) and DNMT1 (Abcam)

2.7.2 Cell culture, Chemicals, Treatments, and Plasmids.

NCCIT and SW480 were obtained from American Type Culture Collection and maintained at 37^{o}C and 5% CO_2 in RPMI-1640 and McCoy's 5a, respectively, supplemented with 10% fetal bovine serum. HCT116 WT, DNMT1 hypomorphic and 3B

KO cells were maintained as described previously (Rhee et al., 2002). DNMT1 hypomorph HCT116 cell line was clonally selected for stable expression of exogenous FLAG-tagged full length DNMT1 and maintained in media containing puromycin. For H_2O_2 exposure, 30% H_2O_2 (Sigma) was diluted in PBS immediately before adding it to the medium. Time of treatment is the time after H_2O_2 is added to the media. c-Myc-Nuc-hOGG1 (Chatterjee et al., 2006) was kindly provided by D. Sidransky (Addgene plasmid 18709). For UV exposure, cells were irradiated with 50 J/m^2 UV using a Stratagene UV crosslinker and incubated at 37 ^oC for 0.5 or 1 hour. For ionizing radiation, cells were exposed to a cesium source at a dose rate of 0.5 Gy/minute for a total of 5 Gy.

2.7.3 Salt Gradient Extraction and Tight Chromatin Fractionation

Cell pellets were collected 30 minutes after H_2O_2 exposure, and sequentially washed in CEBN buffer (10 mM HEPES [pH 7.8], 10 mM KCl, 1.5 mM $MgCl_2$, 0.34 M sucrose, 10% glycerol, 0.2 % NP-40, 1X protease inhibitor cocktail (Invitrogen), 1X phosphatase Inhibitor cocktail and Pefabloc SC AEBSF (Roche Applied Science), N-ethylmaleimide (Sigma), CEB buffer (CEBN buffer without NP-40), soluble nuclear buffer (3 mM EDTA, 0.2 mM EGTA, inhibitors as described above), and salt buffers with increasing NaCl concentration (50 mM Tris pH 8.0, 0.05% NP40, NaCl as indicated, inhibitors as described above). Tight chromatin fractionation was performed essentially the same as the salt gradient except that the only salt wash performed was with 0.45 M NaCl buffer. Western blots of the protein extracts were probed with antibodies listed above. Band densitometry for western blots was analyzed using ImageJ software.

2.7.4 shRNA Knockdown

Cells were transduced with the indicated lentiviral particles following the manufacturer's protocol (Sigma).

2.7.5 Co-Immunoprecipitation

Nuclear extraction was performed using CEBN and CEB buffer as described above. The total nuclear pellet was resuspended in modified RIPA buffer, sonicated for 12 pulses, rotated at 4^{o}C for 30 minutes with 30 mM spermine and 10 mM spermidine to release chromatin bound proteins, sonicated for 10 pulses, and cleared by high-speed centrifugation (20 min, 20,000 x g, 4^{o}C). Nuclear lysates were rotated with antibody overnight at 4^{o}C. Rabbit Trueblot agarose beads (DNMT3B and SIRT1 co-IPs - eBioScience) or Protein A/G-agarose beads (DNMT1 co-IP – SantaCruz BioTechnology) were added and the samples were rotated for 3 H at 4^{o}C. The beads were washed six times with TNE + buffer (50 mM Tris pH7.5, 150 mM NaCl, 5 mM EDTA, 0.5% NP-40, 0.5% TritonX-100, 50 mM NaF, all inhibitors listed above) for 10 minutes at 4^{o}C. For the sequential co-immunoprecipitation, nuclear lysates prepared as above were incubated overnight with FLAG-agarose (Sigma). Complexes were eluted off the beads with FLAG peptide. IPs were then performed from the elutes using anti-EZH2 antibodies.

2.7.6 Sucrose Gradient

Nuclear extracts prepared as for co-immunoprecipitation were applied to a 15-60% (w/v) sucrose gradient, and centrifuged in a SW41 rotor for 20 hours at 40,000 rpm at 4^{o}C. Equivalent volumes from each odd fraction were separated by SDS-PAGE and

analyzed by immunoblot. The remaining fractions were pooled as indicated. Buffer was exchanged to modified RIPA using PD-10 columns (GE Healthcare). Co-immunoprecipitations were performed from each pool as in the section above.

2.7.7 Immunofluorescence

Cells grown on coverslips were preextracted using 0.5% Triton-X 100 in 10 mM HEPES (pH 7.4), 2 mM MgCl2, 100 mM KCl, 1 mM EDTA, fixed in 4% paraformaldehyde, permeabilized in PBS + 0.5% Triton-X 100, blocked in PBST (PBS with 0.1% Tween-20) containing 1% BSA, and incubated with the antibodies indicated above. DAPI was used to stain nuclei. For methanol fixation, cells were fixed in ice cold methanol for 20 minutes.

2.7.8 ChIP-chip

Cells were crosslinked using 1% formaldehyde and 0.5 mM DSG. Nuclear extraction was performed using CEBN and CEB, followed by ChIP-chip as previously described (McGarvey et al., 2008) using antibodies indicated above. Samples were either hybridized to the Agilent 1M custom array for human chromosomes 18, 19 and 21 or the human promoter 244K or 1M ChIP-chip arrays from Agilent Technologies.

2.7.9 Local ChIP

Cells were crosslinked using 1% formaldehyde. Nuclear extraction was performed using CEBN and CEB buffers, followed by ChIP as in (Ohm et al., 2007) using antibodies listed above and primers indicated below.

		Forward Primer	Reverse Primer
ChIP	*MYC*	GCCCTTTCCCCAGCCTTAGC	AACCGCATCCTTGTCCTGTGAGTA
	ACTB	CTCCCTCCTCCTCTTCCTCA	TCGAGCCATAAAAGGCAACTT
	NANOG	TTGTTGCTGGGTTTGTCTTCAGGTT	CGTCTACCAGTCTCACCAAGGCC
	Actb	AGCTTCTTTGCAGCTCCTTCGTTG	CCATGGTGTCCGTTCTGAGTGAT
	Gapdh	CCTGGCACTGCACAAGAAGA	CCACCATCCGGGTTCCTATA
	Sfrp5	AAAGACTCAGGGTTCCTCCCAGGT	TTCGCCGTTCCCTAGCCAATCT
	Sox17	AAATACGTGCCTCAGAGTCTGCCT	TGAAAGGTGCCAATCGACCGCAT
	Fbn1	GGACTGCCTACACGGTCTTAATG	TGGCCATGACCCGGTATG
	Sez6l	AGATCCCTCTGCCTGGATCCCATT	AGAGAGATGGACAGAAGGCTCCAA
RT-PCR	*MYC*	GTCAAGAGGCGAACACACAA	GGCCTTTTCATTGTTTTCCA
	ACTB	GAAGCCGGCCTTGCACAT	AGCACAGAGCCTCGCCTTT
	TIMP3	GGCGGCAGCAGCGGCAATGAC	TACCAGCTTCTTCCCCACCACCTT
	MLH1	GAATGCGCTATGTTCTATTCCATCC	ATAGATCAGGCAGGTTAGCAAGCTG
	NANOG	GCAGAAGGCCTCAGCACCTA	AGGTTCCCAGTCGGGTTCA
	IL8	ACCACCGGAAGGAACCATCTCA	AGCACTCCTTGGCAAAACTGCAC
	HBD	TGACAAGCTGCACGTGGATCC	GGTGAATTCCTTGCCAAAGTTGCG
Bisulfite Sequencing	*MLH1*	TTTTTAGGAGTGAAGGAGGTTA	CCAACCAATAAAAACAAAAATACC
	SFRP5	TTAAATGTTTAGGGAGGTAGGGAGT	AATCGCCCAAATAAATAACAACCTAC
	SFPR4	GAGGGGGTGATGTTATCGTTTTTGTAT	CCCCAAACTCCAATCGACAACAAAAC

2.7.10 Nascent Transcription

Nascent transcription assays were performed using the Click-iT Nascent RNA Capture Kit (Invitrogen). Cells were labeled with ethynyl uridine for 30 minutes concurrently with the H_2O_2 treatment if indicated.

2.7.11 Bisulfite Sequencing

Bisulfite treatment was performed with the EZ DNA Methylation Kit (Zymo). Bisulfite sequencing was performed as previously described (McGarvey et al., 2007).

2.7.12 Mice

C57BL/6J mice were handled and inoculated as in (Rhee et al., 2009). Distal and proximal epithelium was collected by scraping the mucosal surface of the dissected colon, washed three times in PBS, and then subjected to the indicated protocol. Such scraping has been shown by others to be an effective method to obtain samples of intestinal epithelial cells (Ortega-Cava et al., 2006). ChIP from this tissue was performed using the Magna ChIP™ G Tissue Kit (Millipore). All mouse protocols were approved by the Johns Hopkins University Animal Care and Use Committee in accordance with the Association for Assessment and Accreditation of Laboratory Animal Care International.

2.7.13 Statistical Analysis

All western blot, immunofluorescence, and local ChIP data are presented as the mean +/- standard error (SEM). These data are evaluated by one-tail t-test and considered statistically significant with a p-value < 0.05.

ChIP-Chip data were analyzed utilizing the limma and Ringo packages from Bioconductor (Smyth and Speed, 2003; Toedling et al., 2007). ChIP-chip median raw values are normalized through Tukey-biweight scaling procedure and smoothed using an 800 bp sliding window requiring a minimum of 3 probes in the window. ChIP enriched regions are identified as regions with three or more significant probes in row assuming a nonparametric symmetric Null distribution. Box plots are constructed using the ChIP enrichment for probes within the range of 2000 bp upstream and 2000 bp downstream relative to corresponding TSSs. Data were tested using two-tail t-test.

Normalized basal expression array data were obtained from a previous study (Easwaran et al., 2010). Genes with log2 expression levels of greater than 9 or less than 7 are identified as high or low expression genes, respectively. Human transcriptome data were adapted from (Folle et al., 2010). Genomic GC content and CpG island information was extracted from the UCSC genome browser. CpG islands are defined as regions having a GC content of 50% or greater, being longer than 300 bp, and having an observed/expected CG ratio greater than 0.6.

2.8 Accession Numbers

Microarray data sets were deposited in the National Center for Biotechnology Information's Gene Expression Omnibus (GEO) database (http://www.ncbi.nlm.nih.gov/geo) with the accession number GSE32382.

2.9 Acknowledgements

We thank Kathy Bender for manuscript preparation, Dr. D. Sidransky for the c-Myc-nuc-hOGG1 plasmid, and Dr. B. Karim for assessment of H&E stained colon slides.

This work was supported by National Cancer Institute Grant CA043318, National Institute of Environmental Health Sciences grant ES011858, and National Institutes of Health Grant CA116160, all to S.B.B, as well as National Institutes of Health Grants R01DK080817 and R01CA151325 awarded to C.L.S. and National Cancer Institute Grants CA51085 and CA98454 to R.A.C.. C.D.S. is supported by National Institute of Environmental Health Sciences Training grant ES07141.

Chapter 3

A genomic approach for defining potential efficacy of 5-aza-cytidine in non-small cell lung cancer

Revised version published: Alterations of immune response of non-small cell lung cancer with Azacytidine.

Oncotarget. 2013 Nov;4(11):2067-79.

3.1 Abstract

The DNA demethylating agent 5-aza-cytidine (5-Aza-CR) shows therapeutic promise for advanced, non-small cell lung carcinoma. To help understand this, we match genome-wide DNA methylation and gene expression responses to 5-Aza-CR in cell lines with basal levels of these parameters in the Cancer Genome Atlas (TCGA) project. The response signature for the most robust cell line anti-tumor response clusters a small group of TCGA lung squamous tumors, and couples low expression and DNA hypermethylation of *RASSF1* with expression changes, suggesting high cell cycle entry and stem cell pathway activity and low apoptotic capacity. A second response signature, clusters larger groups of TCGA tumors with downregulation of interferon regulated immune responses and the immune tolerance ligand PD-L1. We, thus, suggest a method for deriving predictive biomarkers for those NSCLC patients responding robustly to epigenetic therapy alone, and a larger group for which early clinical data suggest such therapy sensitizes to targeting of immune tolerance.

3.2 Statement of Significance

A pre-clinical, genomics based, approach explores promising responses of patients with non-small cell lung cancer to therapy employing the DNA de-methylating agent 5-AZA-cytidine (5-Aza-CR). Cell line genomic data for responses to 5-Aza-CR are matched to primary tumor data in the Cancer Genome Atlas (TCGA) project. Signatures are derived with relevance to patients responding to the therapy alone, and those responding subsequently to therapy disrupting immune tolerance.

3.3 Introduction

Innovative strategies are desperately needed to treat the world's most common cause of cancer death, non-small cell lung cancer (NSCLC) (Siegel et al., 2013; Youlden et al., 2008). Only 15-20% of lung adenocarcinomas (LUAD) harbor genetic abnormalities that yield high frequency responses when targeted and are generally followed by acquired therapeutic resistance (Shepherd et al., 2005; Vadakara and Borghaei, 2012). In the current study, we take a novel genomics based approach to suggest how targeting epigenetic abnormalities, associated with frequent gains in gene promoter DNA methylation (Baylin, 2011; Jones and Baylin, 2007), may help fill the void for efficacious treatments to manage advanced NSCLC. The use of low doses of the DNA demethylating agents, 5-aza-cytidine (5-Aza-CR) and 5-aza-deoxycytidine (DAC), is the standard of care for patients with certain hematologic neoplasms (Issa and Kantarjian, 2005; Silverman and Mufti, 2005). Our recent laboratory studies reveal that transient exposure of solid tumor cells to low doses of these agents can lead to a reprogramming response which affects populations of tumor self-renewing cells and can blunt their tumorigenicity (Tsai et al., 2012). Thus, 5-Aza-CR and DAC might well be useful therapeutic agents for the most common cancers. Indeed, our recent clinical trial with low dose 5-Aza-CR, plus the histone deacetylase inhibitor (HDACi), Entinostat, shows true therapeutic promise (Juergens et al., 2011) in patients with heavily pre-treated NSCLC of both squamous cell (LUSC) and adenocarcinoma (LUAD) histologies. About 3% of patients experience major RECIST criteria responses which are remarkably durable (Juergens et al., 2011). Interestingly, approximately 20% of NSCLC patients receiving epigenetic therapy had

robust clinical responses to their immediately subsequent therapy. If continued, this would be a very high rate of benefit for the advanced stage NSCLC in the setting of resistance to multiple previous therapies.

The subsequent therapy responses, which have continued to show durability (Figure 3.1) included standard cytotoxic drugs (Juergens et al., 2011), and, and as will be discussed later, a new form of immunotherapy targeting the PD1 immune tolerance checkpoint. While additional clinical trials are needed to validate each of these promising observations, and are now starting, these trials will benefit greatly from insights into potential mechanisms underlying responses. Biomarker strategies to identify the best candidate patients for response to epigenetic therapy alone and those who will receive maximum benefit from subsequent therapies are critical. To begin acquiring such data we have now employed the laboratory model (Tsai et al., 2012) mentioned above with multiple cell lines of NSCLC to explore how low-dose 5-Aza-CR may work for these cancers. We first studied the phenotypic responses to the drug matched with timed genomic analyses after drug withdrawal. We then explored the implications of all the data obtained to primary NSCLC by extrapolating relevant gene expression events and DNA methylation changes to basal DNA methylation and gene expression signatures for hundreds of primary LUSC and LUAD samples (Hammerman et al., 2012) in The Cancer Genome Atlas project (TCGA). The results suggest a genomic strategy for developing drugs, which target epigenetic abnormalities in cancer. Most specifically, they show potential signatures for groups of NSCLC in TCGA involving DNA hypermethylation linked to stem cell, cell cycle, apoptosis, and immune evasion pathways. Our findings suggest key potential mechanisms that may underlie clinical efficacy of 5-AZA-CR for

NSCLC and suggest strategies for biomarker identification to optimize the use of this and related drugs for this deadly tumor type.

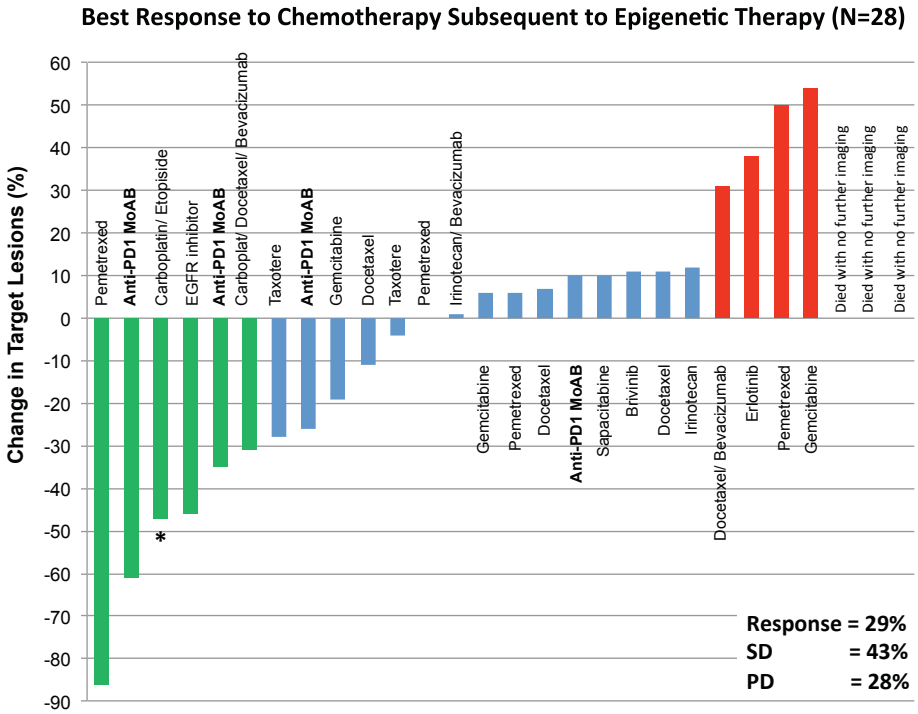


Figure 3.1. Waterfall plot of response to immediate subsequent therapy after progression on epigenetic therapy.

The best change in defined target lesions, as assessed by RECIST criteria scan interpretation, to subsequent systemic anti-cancer treatment following epigenetic therapy is shown. The upward red bars indicate increases in tumor size that constitute disease progression, the light blue bars constitute stable disease, and the green downward bars represent tumor shrinkage qualified to be RECIST criteria response. Green, partial response (PR); blue, stable disease (SD); red, progressive disease (PD). The black text at the bars specifies the type of subsequent therapy the patient received with the bold text showing those who received immune checkpoint immunotherapy. * Indicates a patient who received two rounds of epigenetic therapy. This patient received anti-PD-L1 antibody after his second round of 5-Aza-CR and Entinostat and experienced stable disease.

3.4 Results

We initiated studies using an *ex vivo* model in which cells are transiently exposed to 500 nM 5-AZA-CR. The drug is then withdrawn and the consequences for tumorigenicity are mapped in untreated, immuno-incompetent mice. In turn, prior to explanation, cells are studied for timed, drug-induced, genome-wide changes in DNA methylation and gene expression. We performed such studies on 8 established cell lines of NSCLC, and then explore the implications for primary NSCLC querying how the leading changes map for genome wide patterns of DNA methylation and RNA-seq gene expression analyses in hundreds of tumor samples in TCGA.

For the 8 NSCLC lines, 4 of the 8 lines showed no anti-tumor response, another 2 showed only modest effects, while 2 had extended latency for tumor appearance (Figure 3.2). Of those two, LUAD line (H838), stands out from the other 7 lines by responding even to 100 nM 5-Aza-CR while the second LUAD line H1299 responds only to 500 nM of drug (Figure 3.2). We then used our genomic approach to search for parameters, which correlate with anti-tumor effects (Figure 3.3 & 3.4). Two of the lines which showed no anti-tumor effects, HCC4006 and HCC827, had little to no drop in overall DNA methylation at either the 3-day point of drug administration or 7 days after withdrawal (Figure 3.4A). All other lines exhibited an overall drop in DNA methylation of between 20-40%. The degree of global de-methylation did not correlate with anti-tumor responses (Figure 3.4A).

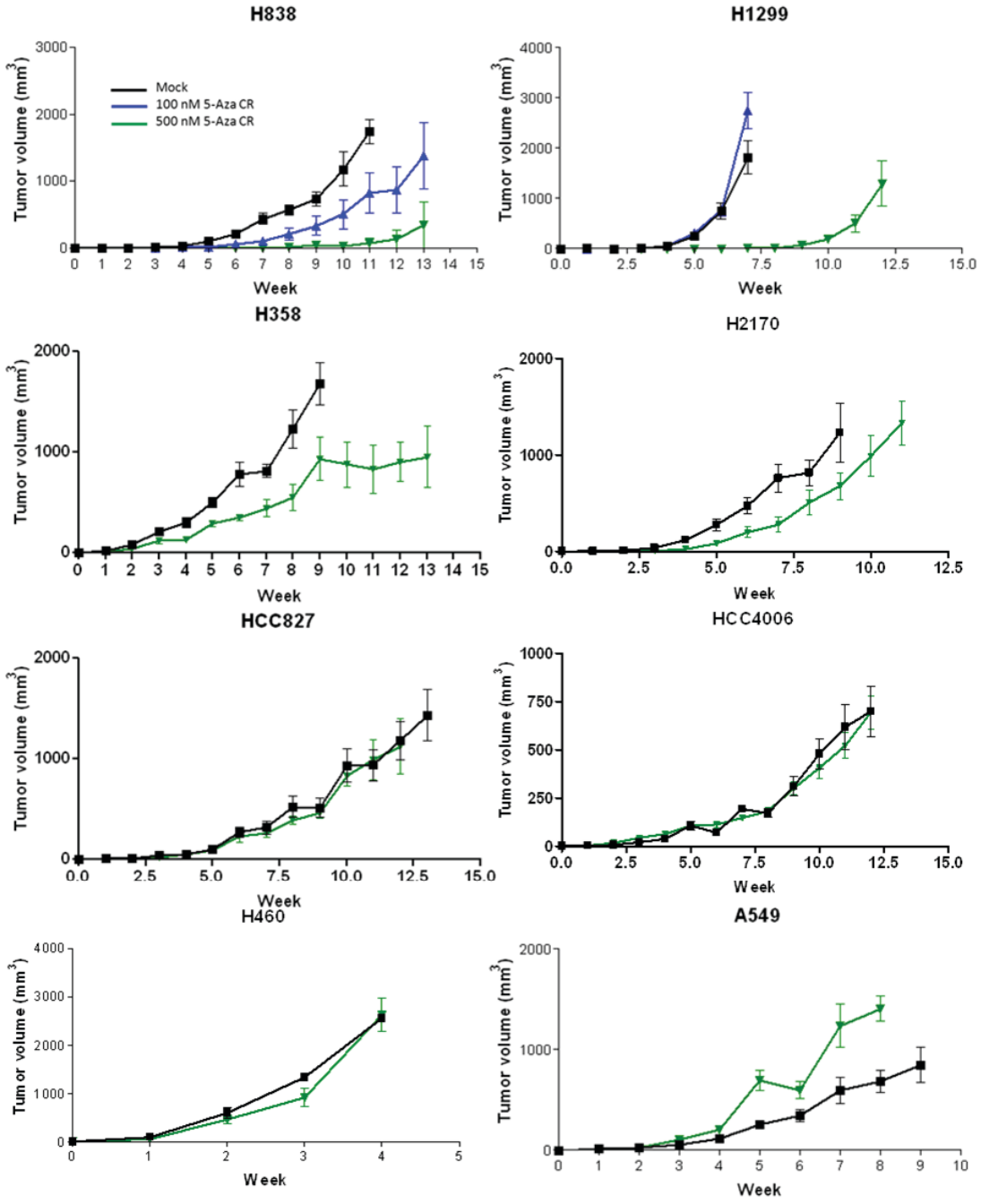


Figure 3.2. Xenografted NSCLC cell lines treated in vitro with 5-Aza-CR in NOD/SCID mice.

Eight NSCLC cell lines were exposed *in vitro* for 72 hr to 100nM and/or 500 nM 5-Aza-CR per a previously published paradigm. Mice were never treated with drug. Tumor volumes are plotted over time. Black lines indicate data for mock treated cells, blue lines for 100 nM 5-Aza-CR treated cells, and green lines for 500 nM 5-Aza-CR treated cells.

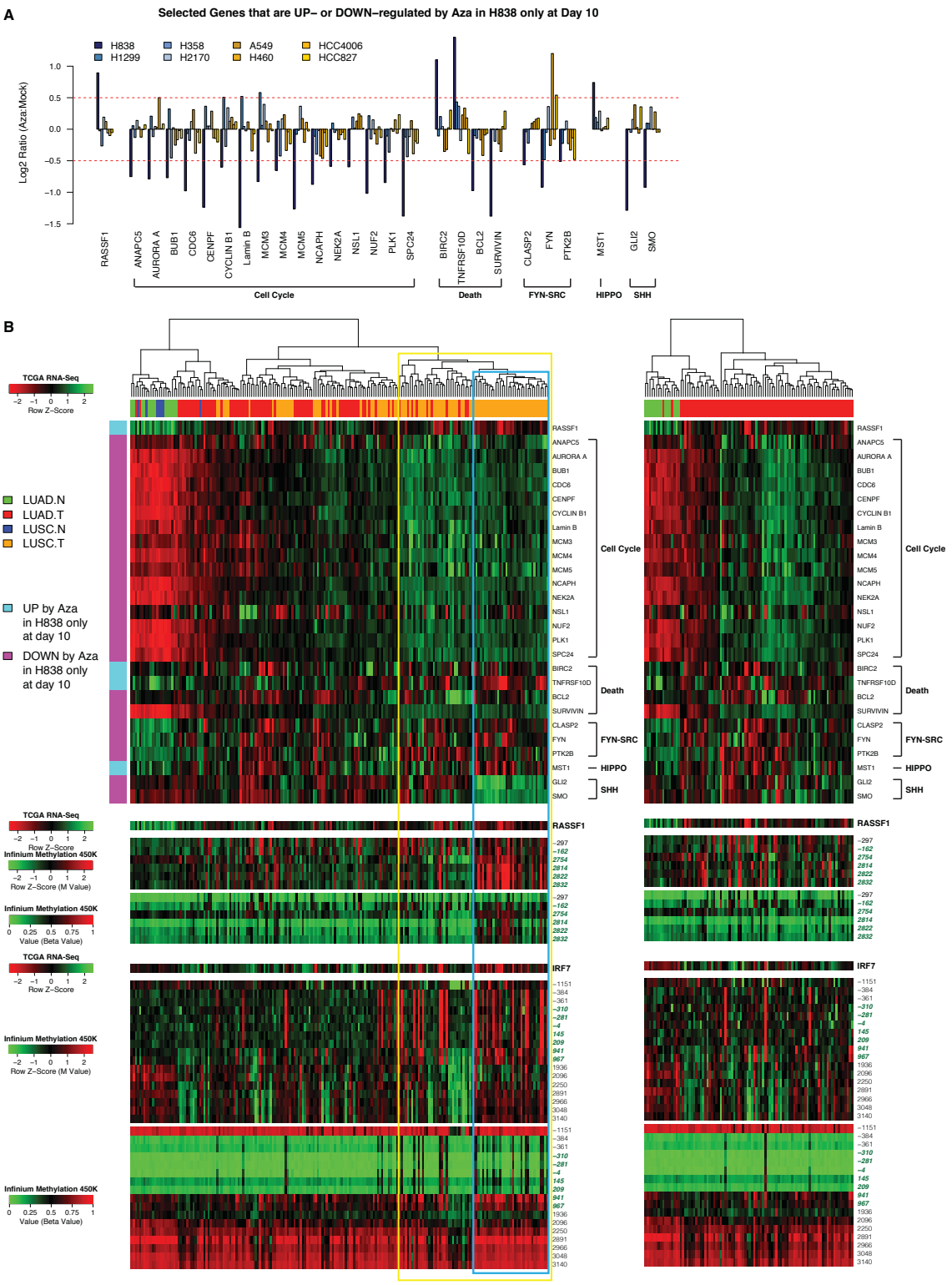

A
Selected Genes that are UP– or DOWN–regulated by Aza in H838 only at Day 10
H838
H1299
H358
H2170
A549
H460
HCC4006
HCC827
Log2 Ratio (Aza:Mock)
1.0
0.5
0.0
−0.5
−1.0
−1.5
RASSF1
ANAPC5
AURORA A
BUB1
CDC6
CENPF
CYCLIN B1
Lamin B
MCM3
MCM4
MCM5
NCAPH
NEK2A
NSL1
NUF2
PLK1
SPC24
BIRC2
TNFRSF10D
BCL2
SURVIVIN
CLASP2
FYN
PTK2B
MST1
GLI2
SMO
Cell Cycle
Death
FYN-SRC
HIPPO
SHH
B
TCGA RNA-Seq
Row Z−Score
LUAD.N
LUAD.T
LUSC.N
LUSC.T
UP by Aza in H838 only at day 10
DOWN by Aza in H838 only at day 10
Infinium Methylation 450K
Row Z−Score (M Value)
Value (Beta Value)
IRF7

Figure 3.3. 5-Aza-CR alters gene expression of multiple pathways related to the *RASSF1* tumor suppressor gene in NSCLC and the relationship to primary NSCLC samples in TCGA.

(A) Effects of 5-Aza-CR mediated expression changes at day 10 in key genes from RASSF1 related pathways outlined in the text. The Y-axis displays the ratio of expression values (log_2) of 5-Aza-CR treated *vs.* mock treated cells for individual genes labeled on the X-axis. (B) Heat map for RNA-Seq expression levels (red for lower and green for higher expression values) for samples of primary lung cancers for genes shown in panel (A). Top bar: red indicates LUAD and orange indicates LUSC with matching adjacent normal lung samples labeled in green and blue as indicated (left panel: all tumor and normal samples; right panel: LUAD alone). Genes used in the heat map are listed on the Y-axis. Bar panels below the main heatmap show the RNA-Seq expression for the genes *RASSF1* and *IRF7* (red for lower and green for higher expression) Below the bar indicating expression for each gene are heat maps of DNA methylation levels across the promoter region probes. Probe positions relative to transcription start site are shown to the right of the heatmap. CpG island probes are labeled in green. Yellow and blue boxes highlight tumors with special features discussed in the text.

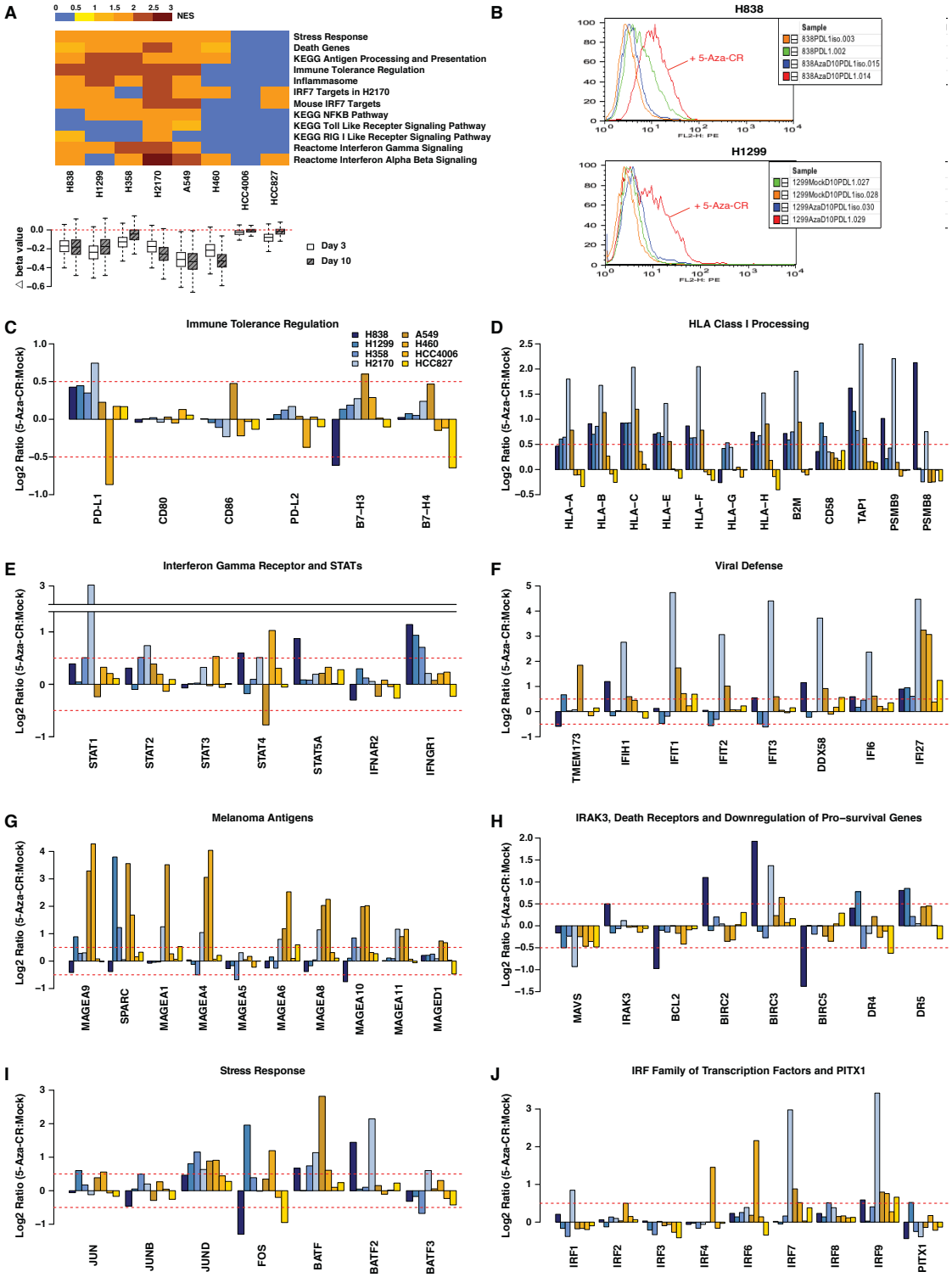

A
NES
Stress Response
Death Genes
KEGG Antigen Processing and Presentation
Immune Tolerance Regulation
Inflammasome
IRF7 Targets in H2170
Mouse IRF7 Targets
KEGG NFKB Pathway
KEGG Toll Like Recepter Signaling Pathway
KEGG RIG I Like Recepter Signaling Pathway
Reactome Interferon Gamma Signaling
Reactome Interferon Alpha Beta Signaling
H838
H1299
H358
H2170
A549
H460
HCC4006
HCC827
beta value
Day 3
Day 10
B
H838
+ 5-Aza-CR
Sample
838PDL1iso.003
838PDL1.002
838AzaD10PDL1iso.015
838AzaD10PDL1.014
H1299
1299MockD10PDL1.027
1299MockD10PDL1iso.028
1299AzaD10PDL1iso.030
1299AzaD10PDL1.029
C
Immune Tolerance Regulation
Log2 Ratio (5-Aza-CR:Mock)
PD-L1
CD80
CD86
PD-L2
B7-H3
B7-H4
D
HLA Class I Processing
HLA-A
HLA-B
HLA-C
HLA-E
HLA-F
HLA-G
HLA-H
B2M
CD58
TAP1
PSMB9
PSMB8
E
Interferon Gamma Receptor and STATs
STAT1
STAT2
STAT3
STAT4
STAT5A
IFNAR2
IFNGR1
F
Viral Defense
TMEM173
IFIH1
IFIT1
IFIT2
IFIT3
DDX58
IFI6
IFI27
G
Melanoma Antigens
MAGEA9
SPARC
MAGEA1
MAGEA4
MAGEA5
MAGEA6
MAGEA8
MAGEA10
MAGEA11
MAGED1
H
IRAK3, Death Receptors and Downregulation of Pro-survival Genes
Log2 Ratio 5-(Aza-CR:Mock)
MAVS
IRAK3
BCL2
BIRC2
BIRC3
BIRC5
DR4
DR5
I
Stress Response
JUN
JUNB
JUND
FOS
BATF
BATF2
BATF3
J
IRF Family of Transcription Factors and PITX1
IRF1
IRF2
IRF3
IRF4
IRF6
IRF7
IRF8
IRF9
PITX1

Figure 3.4. 5-Aza-CR alters gene expression in NSCLC for multiple immune-related pathways and the degree of change is associated with degree demethylation.
(A) Top panel: Gene Set Enrichment Analysis (GSEA) showing upregulation of immune pathway genes by 5-Aza-CR in 8 NSCLC cell lines 7 or 10 days after drug withdrawal. Normalized enrichment scores are plotted as a heat map. Bottom panel: boxplots showing different degrees of demethylation in each of the 8 cell lines as measured by the difference in beta values between the 5-Aza-CR and mock treated cells. (B) FACS analysis showing increased cell surface PD-L1 after 5-Aza-CR treatment at day 10 in H838 and H1299. (C) to (J) Effects of 5-Aza-CR mediated expression changes at day 10 in key genes from the pathways represented in (A). Y-axis displays the ratio of expression values (log_2) of 5-Aza-CR treated *vs.* mock treated cells; X-axis displays names of individual genes.

For our most sensitive cell line, H838, we noted that at the time of cessation of drug treatment on day 3 as well as 7 days later, 5 of the top 10 pathways identified for 5-Aza-CR induced expression changes for genes involved with the ability of the cell to enter and progress through the cell cycle (p-value $<10^{-4}$ to $<10^{-8}$, data not shown). These pathway analyses also linked these changes to up-regulation of *RASSF1*, one of the most frequently hypermethylated tumor suppressor genes in lung cancer and other malignancies (Dammann et al., 2000) (Agathanggelou et al., 2001). Indeed, while 5 of the 8 lines tested harbor DNA hypermethylation across the promoter region CpG island of *RASSF1*, and have some drug-induced decrease in DNA methylation in the region (Figure 3.5), only H838 cells up-regulate the gene transcript (Figure 3.4A). For the above genes noted in Metacore, the link to RASSF1 reflects the fact that this tumor suppressor has solid molecular relationships to suppression of cell mitosis through down-regulation of events in the PLK1, CDK1, Cyclin B, APC^{Cdc20}, and FOXM1 pathway control for entry and progression through the cell cycle (Amin and Banerjee, 2012; Chow et al., 2012; Song et al., 2004). Notably, RASSF1 function is also tied to triggering apoptosis though enhancement of TRAIL-Death receptor (DR) signaling (Amin and Banerjee, 2012) and the conversion of the HIPPO-YAP pathway from oncogenic to anti-tumor effects (Pan, 2010). Finally, its loss of function parallels increases in the sonic hedgehog stem cell pathway in basal cell carcinomas (Brinkhuizen et al., 2012). Remarkably, and virtually uniquely in the H838 cells, there are coordinate changes in expression for genes in all of the above pathways. Thus, there are 5-AZA-CR induced decreases in multiple genes for entry into and progression through, the cell cycle, many tied to the FOXM1 pathway, in concert with the *RASSF1* transcript increases (Figure 3.3A). These events

include 2-fold down-regulation of the key oncogene, *E2F1* in H838 cells, while only one other cell line, H460 (data not shown), has a marginal decrease. This protein interacts with the RB tumor suppressor to increase cell cycle activation (Dimova and Dyson, 2005). These cell cycle gene changes were accompanied in H838 cells by 5-Aza-CR induced decreases for the pro-survival genes *BCL2* and *BIRC5* (*SURVIVIN*) and increases in the pro-apoptosis genes TRAIL receptors, *TNFRSF10B* and *TNFRSF10D* as well as the autophagy genes *BIRC2* and *BIRC3* (Figure 3.3A and data not shown). Finally, the key driver genes for the SHH pathway, *SMO, GLI1*, and *GLI2*, were all decreased and there was concordant down-regulation of the *FYN* oncogene and many related members of the SRC related protein-tyrosine kinase pathway (Montero et al., 2011), and linked to *ITGB4*, a key integrin related to metastatic lung cancer cell behavior (Yang et al., 2010) (Figure 3.3A and data not shown).

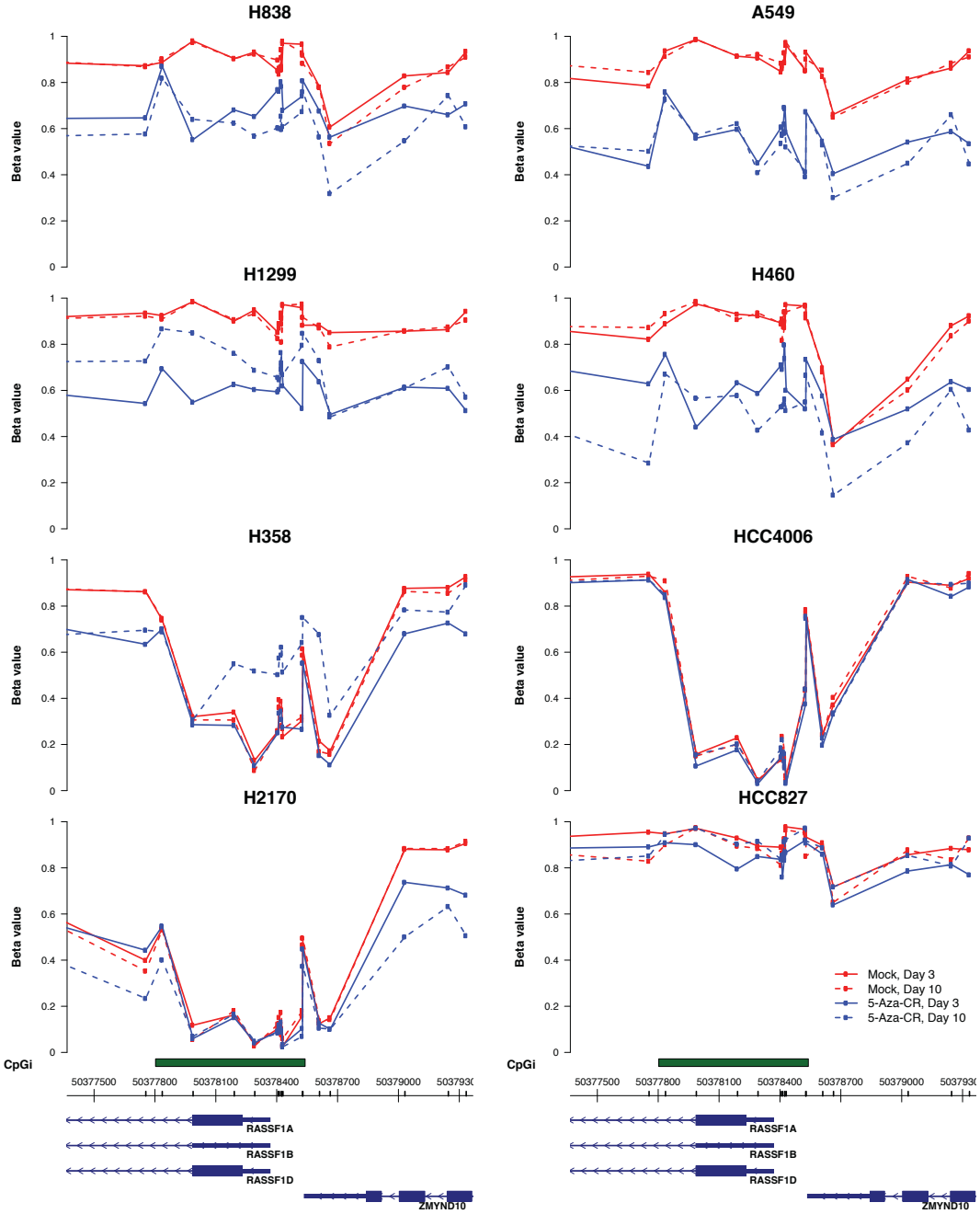


Figure 3.5. Cell line promoter DNA methylation status, assessed as beta value on the Y-axis obtained from Illumina Infinium HumanMethylation450 BeadChip. Genomic location is labeled on the X-axis with black vertical bristles indicating the position of probes. Gene structures are displayed at the bottom. Green box indicates the location of CpG island. Red lines indicate data from mock treated cells, and blue lines from 5-Aza-CR treated cells. Solid lines indicate data at day 3, and dashed lines at day 10.

We next examined how 5-Aza-CR induced gene expression changes unique to H838 cells might relate to hundreds of primary LUAD and LUSC TCGA samples (Hammerman et al., 2012). We examined the promoter CpG island DNA methylation status of *RASSF1* as analyzed by the identical Infinium 450K used in the above cell culture studies, and expression of this gene and other gene changes unique to H838 cells, as mapped by RNA-seq, as outlined in Figure 3.3A. Two important signatures emerged which clustered tumors from one another and separated them from normal lung samples in TCGA. First, a group of tumors (within the yellow box, Figure 3.3B), virtually all LUSC, and constituting 36% of the 195 NSCLC samples, have higher basal expression of genes for control of cell cycle entry and progression. Second, within this group there is a smaller subgroup (blue box, Figure 3.3B, 18% of NSCLC cases) wherein *RASSF1* hypermethylation and lower expression accompany not only higher basal expression of the above genes for control of cell cycle entry and progression, but also of key SHH activation genes, and of the pro-survival genes *SURVIVIN*, and, in a smaller group, *BCL2*. In concert, basal levels are decreased for pro-apoptotic genes, especially notable for *TNFRSF10D* (Figure 3.3B). Third, when LUAD is depicted separately (Figure 3.3B, right panel) it can be appreciated that there is a group of tumors having some of the basal gene patterns for all the pathways as outlined for LUSC, but especially for higher expression of all the cell cycle genes. Thus, the above smallest groups of LUSC and LUAD patients could relate to patients likely to benefit from 5-Aza-CR alone. Indeed, only a small group of patients with advanced NSCLC showed RECIST criteria responses to epigenetic therapy alone, but these were extremely durable. Moreover, both groups of LUAD and LUSC driven by the cell cycle gene pattern may also relate to patients who subsequently

respond to cytotoxic therapies as promising data indicate in the clinical trial (Figure 3.1). For example, inhibitors of FOXM1, PLK1, CDK's, CHEK1, and AURORA kinases are all being developed with the goal of re-establishing cell cycle checkpoints to sensitize patients to chemotherapies (Campos and Dizon, 2012).

We followed another lead from the Metacore analysis of gene expression responses to 5-Aza-CR and which treatment which related to multiple of the cell lines. Multiple immune –related pathways were among the top-most altered in response to 5-Aza-CR (p-value $<10^{-3}$ to $<10^{-5}$, data not shown) in virtually all the treated cell lines. These changes indicate an up-regulation by 5AZA-CR for a complex tumor immune evasion signature. Given the early suggestion in our epigenetic therapy trial that the regimen may provide subsequent sensitization to therapy targeting immune tolerance checkpoints, we intensely pursued these findings. The Gene Set Enrichment Analysis showed significant up-regulation of immune-related pathways in a manner roughly correlating to the degree of demethylation in response to 5-Aza-CR treatment (Figure 3.4A). 5-Aza-CR has been well described by others to up-regulate individual steps in immune pathways including those for tolerance-maintaining proteins in lymphocytes, assembly of major histocompatibility antigens (MHC or HLA Class I), interferon pathway genes, and cancer-testis and other surface antigens (Claus et al., 2005; Fonsatti et al., 2007; Simova et al., 2011). Each of these events is required for the immune tolerance pathways targeted by immune checkpoint therapy (Pardoll, 2012) and some of these genes have low expression in cancers associated with cancer-specific increases in promoter region DNA methylation, and increased expression after treatment with DNA demethylating drugs (Kulaeva et al., 2003; Li and Tainsky, 2011). Importantly, when compared to normal bronchial epithelial

cells, NSCLC cells are known to exhibit diminished innate immune responses to viral-like stimuli, involving intertwined pathways of cell-intrinsic responses to infection and inflammation (Li and Tainsky, 2011). Finally, NSCLC commonly express low levels of genes for antigen processing and presentation, which are required for recognition by tumor-targeting T-cells (Li and Tainsky, 2011). Given all of the above, we extensively mined our DNA methylation and gene expression data for how our pre-clinical model of low dose 5-Aza-CR might alter immune responses.

As might be expected, given that our tumor responses are monitored in immuno-incompetent mice, the immune pathway responses do not correlate with extent of anti-tumor response, as do the pathways analyzed in our studies in Figure 3.2 and 3.3. However, as shown in schematic form (Figure 3.6), and in GSEA (Figure 3.4A), and in examination of individual genes (Figure 3.4C-J), low-dose 5-Aza-CR, at day 10, induces an extensive, complex reprogramming of the multiple immune pathways mentioned. A first key point is that activation of the JUN/FOS signaling response to cell stress (Hernandez et al., 2008; Van der Velden et al., 2012) is central to activation of all of the above immune-related pathways (Figure 3.4A & 3.6), and treatment with 5-Aza-CR up-regulates expression of multiple central member genes including *JUN, JUNB, JUND, FOS*, and key interacting factors *BATF, BATF2*, and *BATF3* (Figure 3I). Members of the BATF family are emerging as particularly important transcriptional regulators of immune cell differentiation and function (Glasmacher et al., 2012; Logan et al., 2012). Next, there is increased expression of a series of genes critical for initial recognition and processing of viral RNA and subsequent triggering of interferon pathways (Figure 3.4A, 3.6) including *STING* (*TMEM173*), *IFIH1* and interacting factors *IFIT1*, *IFIT2* and *IFIT3*, and

IFITM1, *IFITM2*, *IFITM3*, and *IFITM4* and *RIG I* (*DDX58*) (Hsu et al., 2012; Ishikawa et al., 2009) (Figure 3.4F and data not shown). In parallel, genes for initiation and maintenance of the interferon response to viral infection, such as *G1P3* (*IFI6*), are concordantly up-regulated (Figure 3.4F) as are many downstream pro-inflammatory pathway genes linked to viral defense responses (Cheriyath et al., 2011) including multiple interferon, TNF-alpha, and cytokine genes.

Importantly, we observe multiple clues for how agents like 5-Aza-CR might sensitize NSCLC to anti-immune tolerance therapy. One key component of immune tolerance, and effective for inhibiting it, is the dependence of both processes on recognition by T-cells of surface antigens, both of a general nature, and very specifically of MHC or HLA Class I targets (Pardoll, 2012). In terms of the former, as recognized by others (Claus et al., 2005; Fonsatti et al., 2007; Simova et al., 2011), 5-Aza-CR increases expression of multiple cancer testes antigens (Figure 3.4G) including multiple MAGE family genes, whose expression has been shown to be suppressed in non-transformed cells via promoter hypermethylation (Claus et al., 2005; Fonsatti et al., 2007). Also, 5-Aza-CR up-regulates not only transcripts for virtually all HLA Class I antigens but also, concordantly, of a series of genes including, beta-2-microglobulin (*B2M*), *TAP1*, and the immuno-proteasome subunits *LMP-2* (*PMSB9*) and *LMP-7* (*PSMB8*) (Figure 3.4D), which encode proteins required for endoplasmic reticulum processing, transport, and anchoring of MHC-I subunits to the cell surface, and *CD58*, which helps MHC-I mediate immune cell attraction (Challa-Malladi et al., 2011; Procko and Gaudet, 2009; Raghavan et al., 2008). Of note, there are mutations of *HLA-A* in a small percentage of LUSC (Hammerman et al., 2012) and also of *B2M* and *CD58* in other tumor types (Challa-Malladi et al., 2011).

These mutations have all been incriminated as genetic events which can provide mechanisms for immune evasion.

Second, *in vivo*, 5-Aza-CR administration to tumor-bearing mice has been shown to induce antigen processing and presentation genes, particularly when administered with CpG TLR9 agonists, and this is largely attributed to interferon-γ production by lymphocytes (Simova et al., 2011). While the lymphocyte-specific γ-interferon is not ectopically induced in NSCLC lines with 5-Aza-CR treatment, there is robust induction of the gene encoding the inducible subunit of the interferon-γ receptor (*IFNGR1*), but not the interferon-β receptor (*IFNAR2*), as well as of multiple STAT genes, including *STAT1*, the major *IFNGR1* signal transducer (Figure 3.4E).

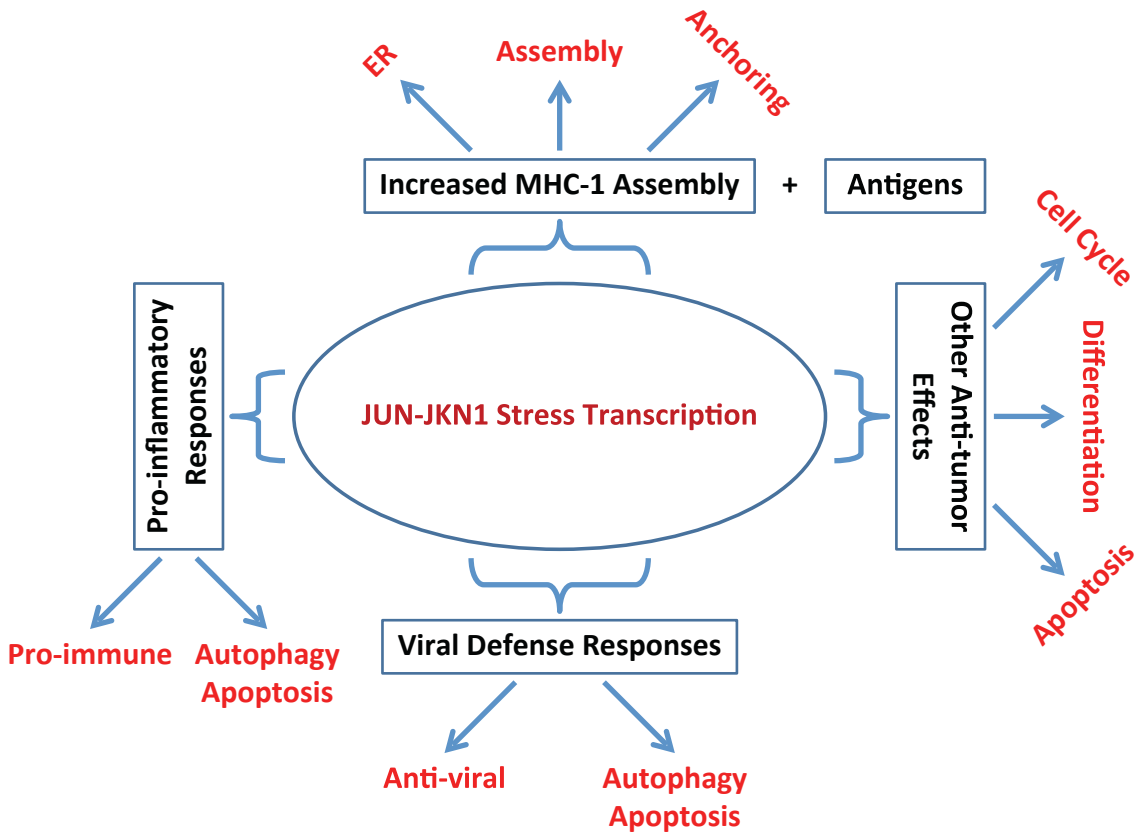


Figure 3.6. Schematic of the immune and death related processes that were analyzed for genes curated from the literature for key pathways.

Importantly, for the cell studies, 5-Aza-CR treatment influences a key tumor cell event, presence of PD-L1 on the cell surface, which is an essential target for therapy to inhibit immune tolerance. This is the major immune checkpoint ligand for the interacting immune cell receptor, PD-1 (Pardoll, 2012). In the reported anti-PD-L1/PD-1 clinical trials (Brahmer et al., 2012; Topalian et al., 2012), in a small subset of the patients analyzed, responses did not occur in patients whose tumors did not express cell surface PD-L1 (Topalian et al., 2012). 5-Aza-CR increases the *PD-L1* much more frequently than it does for *PD-L2*, a second dendritic cell/macrophage ligand for PD-1, or other checkpoint ligands such as *B7-H3* and *B7-H4* (Pardoll, 2012) (Figure 3.4C). Importantly, *CD80* and *CD86*, ligands for the immune checkpoint receptor CTLA4, another target in immunotherapy strategies (Pardoll, 2012) are not altered (Figure 3.4C). Importantly, *PD-L1* expression in tumor cells can either be driven by cell-intrinsic mechanisms, or in a process termed adaptive resistance (Pardoll, 2012), by interferon-γ signaling and subsequent activation of STAT transcription factors, which we show earlier to be induced by 5-Aza-CR (Figure 3.4E). We have performed FACS analyses for expression of PD-L1 protein on the surface of two of the NSCLC lines (Figure 3.4B). Consistent with the transcript changes, 5-Aza-CR increases this tumor ligand, as compared to mock treatment in both instances. We also looked at other selected immune-related proteins HLA Class I, B2M, CD58, and B7-H3 for their appearance on the cell surface and found generally good correlation with 5-Aza-CR induced transcript increases (Figure 3.7).

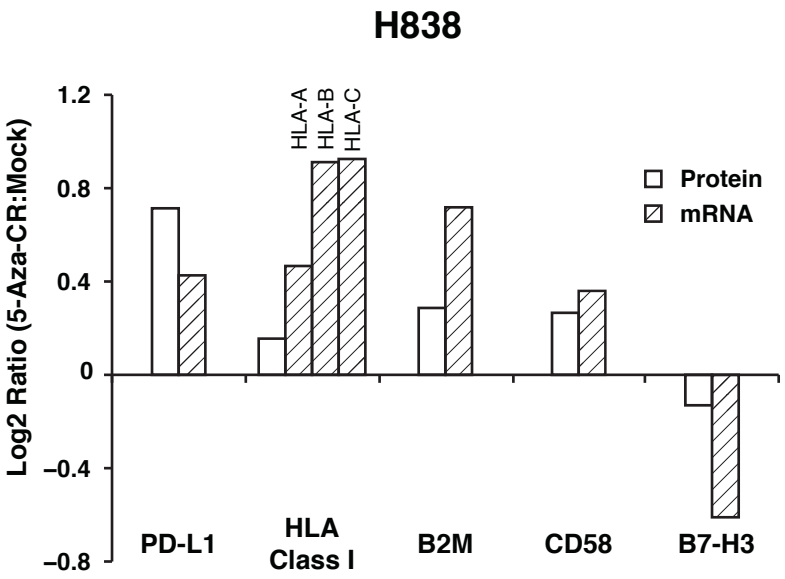


Figure 3.7. Comparison of Agilent expression array data to flow cytometry for select cell surface proteins in H838.

Clear bars represent the $\log_2$ ratio of mean fluorescence intensity of 5-Aza-CR over mock treated cells. Hashed bars represent the M-values of expression array ($\log_2$[5-Aza-CR:Mock]). For HLA Class I, the antibody used in flow cytometry does not discriminate subtypes of class I molecules. Individual class I molecule subtype transcript data are available from the Agilent array platform and is presented.

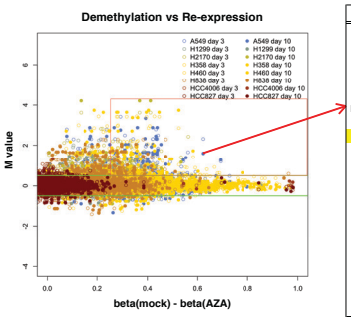
A final key point to be addressed for all the cell line data above is a potential intersection between the immune pathway events and the other pathways which have emerged, especially in triggering cell death. These overlaps may be important determinants of the anti-tumor responses in our most sensitive cell lines, even in an immuno-incompetent mouse, and in patient responses. In this regard, the pro-inflammatory and viral defense pathways harbor what has been termed "the dual face" of pro-inflammatory and viral defense responses for promoting or inhibiting cell death and/or tumorigenesis (Dunn et al., 2002; Ishikawa et al., 2009; Sharma and Fitzgerald, 2010). These death signals are triggered predominantly by increased extent and duration of cell challenges, like viral infection, such that involved cells must be eliminated (Dunn et al., 2002; Ishikawa et al., 2009; Sharma and Fitzgerald, 2010). Many of the apoptotic pathways involved trigger the same down-stream gene expression events shown in Figure 3.3A and 3.4F, 3.4H. In terms of the specific immune genes with potential for inhibiting tumor growth, 5-Aza-CR triggers up-regulation of *IFI27*, especially in H2170 LUSC cells (Figure 3.4F), which encodes a protein triggering apoptosis in late stages of chronic viral infection (Cheriyath et al., 2011). Simultaneously, there is down-regulation of the anti-apoptotic gene, *MAVS* (Figure 3.4H), a change that accompanies RIG I signaling in response to viral challenge (Sharma and Fitzgerald, 2010; Xu et al., 2010). Downstream events in these viral response death signals involve the signature seen in our most sensitive H838 cells, as previously discussed, including increases for expression of BIRC family, autophagy genes (Yang and Klionsky, 2010) and simultaneous decreases in the anti-apoptotic genes *BCL2* and *BIRC5* (*SURVIVIN*) (Figure 3.4A & 3.4H). Interestingly, the *SURVIVIN* response is known to be triggered by the viral induction of *IRAK3*, a gene

expression increase unique for the H838 cells. This gene encodes for an IL-1 receptor associated kinase (De Carvalho et al., 2012). *IRAK3* is silenced in colon cancer cells in association with promoter-region DNA hypermethylation and when reactivated by demethylation, is associated with *BIRC5* down-regulation (De Carvalho et al., 2012). In addition, many other apoptosis triggering genes central to inflammation and viral-defense (Benedict et al., 2002; Strowig et al., 2012) are altered in our cells including caspases, TNF pathway genes, and interleukins.

To search for potential driving events for all of the immune-related changes induced by the transient 5-AZA-CR treatment, we next examined induced promoter de-methylation and increased gene expression of key upstream transcription factors (TFs) (Figure 3.8A). We identified approximately 300 genes with high baseline promoter region CpG island methylation, promoter demethylation of 25% or more after treatment, and expression increased by $\log_2$ 0.5 (1.4-fold) or greater after treatment (Figure 3.4A, Table 3.1). Nearly 17% of these are in an interferome database (http://www.interferome.org), and 52 (19%) are TFs (Bidwell et al., 2012; Samarajiwa et al., 2009). The TF, *IRF7*, has been reported by others to be hypermethylated in cancer (Jee et al., 2009; Li and Tainsky, 2011; Lu et al., 2000), as it is in our NSCLC line with the lowest basal expression (Figure 3.9). It is up-regulated in response to 5-Aza-CR in several cell lines, most prominently in the LUSC cell line H2170 showing a 9-fold increase (Figure 3.4J) (Li and Tainsky, 2011). This major upstream activator functions in cellular pathways recognizing the virus response element VRE-A to increase transcription of genes involved in type 1 IFN signaling (Li and Tainsky, 2011). PScan analysis confirms, by presence of these sequences, a significant association of IRF7

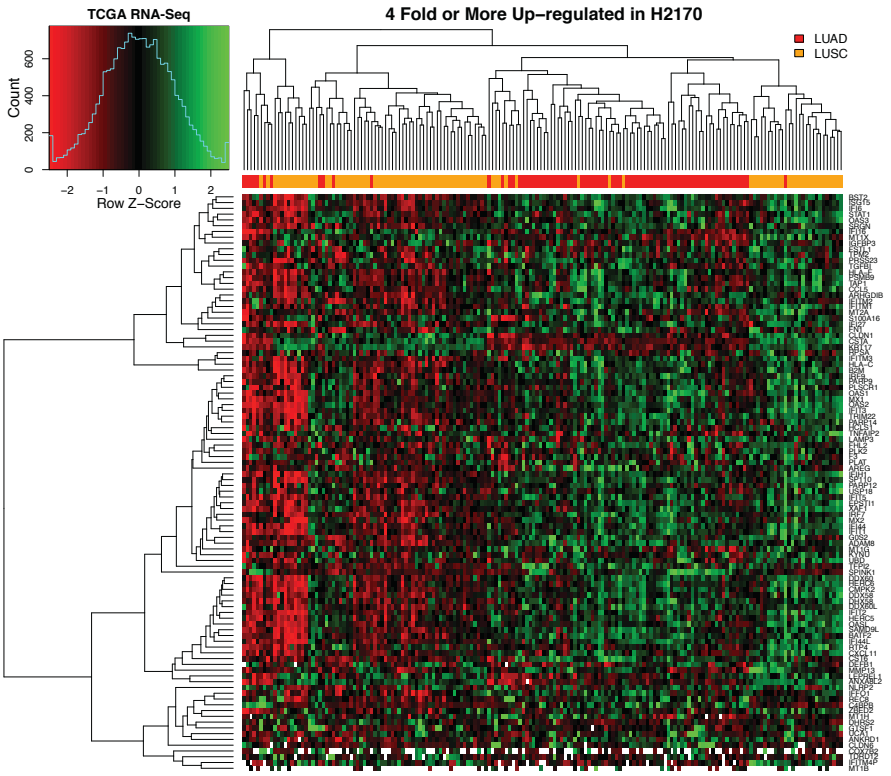
transcription targets with genes which drive GSEA enrichment scores for the immune pathway alterations observed in response to 5-Aza-CR (Table 3.2). The same dynamics hold true for *PITX1*, up-regulated by 5-Aza-CR in our second most anti-tumor responsive cell line, H1299 (Figure 3.4J & 3.9), which acts at the distal negative regulatory element (DNRE) to suppress the activity of IRF family TFs (Bidwell et al., 2012).

A



A549	H1299	H2170	H358	H460	H838	HCC4006	HCC827
ANKRD53	H1F0	ELF4	HIC1	BNC1	HES5		
CREB3L1	HIC1	GATA3	HR	CTCFL	HIST1H2BJ		
CTCFL	HOXD13	HOXC9	MEIS2	DZIP1	IRF7		
ELL3	ID4	IRF7	PDLIM4	ELF4	IRX5		
GATA3	PITX1	KLF11	PITX2	FOXF1	MAFG		
HIST1H1B	TRIP6	PLEK2	PLEK2	H1F0	MST1R		
HIST1H2BG	ZNF239	SOHLH2	TCF15	ID4	SOX15		
HOXA5		ZNF502		IRF6	ZNF300		
IRF7		ZNF578		IRF7			
MEIS2				IRX5			
MST1R				PER3			
PITX2				PLAGL1			
PLEK2				PLEK2			
POU4F1				RUNX3			
TCF15				SOHLH2			
TCF21				SOX11			
TLX2				TCF15			
TRIM29				ZNF167			
ZIK1				ZNF300			
ZNF578				ZNF331			
ZNF75A				ZNF502			
				ZSCAN18			

B



C

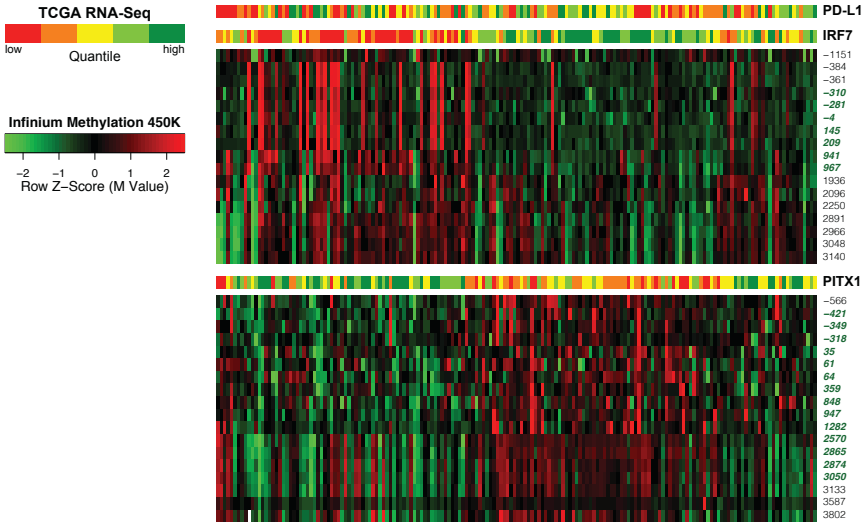


Figure 3.8. Genes highly up-regulated in H2170 by 5-Aza-CR are associated with IRF7 in NSCLC.

(A) Left panel: Plot of genes with high baseline DNA-methylation of promoter region CpG island probes (Infinium beta values > 0.5) which demethylate by 0.25 or more at day 3 and/or day 10 with 5-Aza-CR treatment and which show a gene expression increases of 0.5 or more (log_2) in NSCLC cell lines (color of dots indicate cell line and time point, X-axis indicates differences in beta values between mock treated and 5-Aza-CR treated cells, Y-axis displays expression values as log_2 differences between 5-Aza-CR *vs.* mock treated cells). The red box encompasses genes meeting the described criteria for DNA methylation decrease and gene expression increase. Right panel: Selected transcription factors meeting the described criteria. The full list of genes up-regulated by 5-Aza-CR according to the described criteria is listed in (Table 3.1). Two key immune-related TF genes, *IRF7* and *PITX1*, are highlighted in yellow in the right panel. (B) Heat map of genes derived from the LUSC cell line H2170 using RNA-seq data for primary LUSC and LUAD samples from the TCGA. The genes are those 4-fold or more up-regulated in H2170 in response to 5-Aza-CR in coordination with a 9-fold increase in *IRF7* expression for samples of primary lung cancers (red for lower and green for higher expression values). This gene list is highly associated with IRF7 as transcription factor targets based on PScan analysis (p = 7.6 e -18, Table 3.2). Top bar: red indicates LUAD and orange indicates LUSC. Genes used in the heat map are listed on the Y-axis as well as in Table 3.3. (C) Bar panels show the expression spectrum for the indicated genes (*PD-L1*, *IRF7*, and *PITX1*) in five quantile intervals (red for lower and green for higher expression). Heat maps immediately below each expression spectrum bar panel are Infinium DNA methylation data across the promoter region for CpG probes with positions relative to transcription start site shown to the right (red for more and green for less methylated). CpG island probes are labeled in green. Order of samples in bar panels and heat maps is maintained from hierarchical clustering in panel (B).

Table 3.1. Complete list of genes that are re-expressed for 0.5-fold or more (log_2 scale), promoter region CpG island hypermethylated in mock treated cell, and demethylated for greater than 25% by 5-Aza-CR in the eight NSCLC lines. Identified transcription factors are labeled in red. Genes identified by the interferome database are underlined.

A549	H1299	H2170	H358	H460	H838	HCC4006	HCC827
AJAP1	AKAP12	AKAP12	ADAM23	ACTN2	ALDH1A3		UBE2DNL
AKAP12	ANKK1	BIK	CACHD1	AJAP1	C1orf190		
ALDH1A3	ANPEP	BNIP3	CCK	AKAP12	C2orf84		
AMPD3	APH1B	CD38	COL7A1	AKR1E2	CD248		
ANKK1	BIK	CLDN23	CRHR1	ALDH1A3	CDK5R1		
ANKRD53	C20orf46	COCH	CYB5R2	ALPL	CPT1C		
ANPEP	CLIC3	DDX43	DSE	APH1B	CYP11A1		
BIK	COL9A3	EIF5A2	ENG	AQP5	DNALI1		
C5orf58	DNALI1	ELF4	FBXO2	ASPRV1	EFHD1		
C9orf140	FAM132A	FBXO2	FZD7	BNC1	ELMO3		
C9orf4	FERMT3	FERMT3	HIC1	C20orf46	ESRP2		
CCDC144NL	FMR1NB	FLNC	HIST1H2AD	C8orf84	FAM50B		
CCK	FZD7	FMR1NB	HIST1H2AE	CAND2	FES		
CES1	GDF15	FSTL1	HR	CAPS	GDF15		
CHTF18	GPRC5B	FTHL17	LIMS2	CCDC144NL	HES5		
CLDN4	H1F0	GATA3	MEIS2	CCND2	HIST1H2BJ		
CLIC3	HIC1	HERC5	MFI2	CD248	HSD17B8		
CLIC6	HOXD13	HLA-F	MSI1	CD38	ICAM4		
COL9A3	ICAM1	HOXC9	PCDHGA7	COL9A3	IFFO1		
CREB3L1	ICAM4	HSPA2	PDLIM4	COX7A1	INA		
CTCFL	ID4	IFFO1	PITX2	CPT1C	IRF7		
CYB5R2	LOC151534	IGF2BP1	PLEK2	CTCFL	IRX5		
DDX43	LPIN3	IRF7	PTHLH	D4S234E	ISG15		
DNALI1	LRFN4	IRS1	PTPRS	DBNDD2	KLHDC7B		
DPEP3	LY6K	KLF11	RGMA	DDX43	LOC654433		
DUSP2	MEST	KLHDC7B	SDHAP3	DKK3	MAFG		
ELL3	MFI2	LOC654433	SGK1	DNAJA4	MEI1		
FAM20A	PITX1	MARVELD1	SLC16A1	DNALI1	MST1R		
FAM83H	PLD6	MFI2	TCF15	DPEP3	NEFH		
FBLN2	PLLP	MT1L	TNFRSF25	DZIP1	NES		
FBXO2	PP14571	NES	TTTY14	EHD3	NTF3		
FERMT3	S1PR4	NKAPL	UCHL1	ELF4	POMC		
FES	SLC4A11	PCDHGA3		ELOVL2	PRDX2		
FMR1NB	SPINT2	PLBD1		F2RL1	PSMB8		
FRZB	TMEM204	PLEK2		FABP5	REC8		
FSTL1	TRIM74	PNLDC1		FAM118A	ROBO3		
FZD10	TRIP6	PPP1R14A		FAM150B	S1PR4		
GATA3	TYRO3	REC8		FAM20A	SLC15A3		
GJB2	VWCE	SCIN		FAM3B	SOX15		
GPRASP1	ZNF239	SLC15A3		FAM50B	STAG3		
GSTM2		SLC16A5		FBLN2	TDRD12		
H19		SOHLH2		FBXO2	TMEM88		
HIST1H1B		STOM		FKBP1B	TNFRSF25		

A549	H1299	H2170	H358	H460	H838	HCC4006	HCC827
HIST1H2AD		TDRD12		FMR1NB	VAMP5		
HIST1H2AE		TDRD9		FOXF1	VWCE		
HIST1H2BG		TSPYL6		FRZB	WDR69		
HIST1H3D		VWCE		FZD10	ZNF300		
HLA-F		ZNF502		GALNT3			
HOXA5		ZNF578		GCC1			
HPDL				GDF15			
HRASLS5				GFRA1			
HSPA2				GPC2			
ICAM1				GPX7			
IFFO1				H19			
INPP5D				H1F0			
IRF7				HS6ST1			
ISG15				HSPA2			
ISYNA1				ID4			
JAM3				IFFO1			
KANK4				INPP5D			
LOC654433				IRF6			
LRFN4				IRF7			
LY6K				IRX5			
MAGEB1				ISG15			
MEG3				ISYNA1			
MEIS2				KIAA1614			
MFI2				KLHDC7B			
MST1R				KRT7			
MX1				L1TD1			
MYL9				LAYN			
NAV1				LRRK1			
NEFH				MAGEB1			
NEFM				MEG3			
NINL				MEST			
NUDT11				MEX3A			
PITX2				MFI2			
PLBD1				MT1M			
PLEK2				MX1			
PLLP				MYO5C			
POU4F1				NAPRT1			
PRSS21				NEFH			
RAB31				NINL			
RBM46				NKAPL			
REC8				PCDHGA2			
SEMA6C				PCDHGA3			
SIPA1L3				PCSK9			
SLC12A8				PDIA2			
SLC4A11				PER3			
SPOCK2				PLAGL1			
ST3GAL2				PLBD1			
STAG3				PLEK2			
TCF15				PNLDC1			
TCF21				PNPLA3			
TDRD12				POMC			

A549	H1299	H2170	H358	H460	H838	HCC4006	HCC827
TDRD9				PTGER4			
TLX2				PTHLH			
TMEM204				RASL11B			
TMEM220				RFTN1			
TNFRSF10C				RGAG4			
TNFRSF25				RGMA			
TNPO2				RIBC2			
TRIM29				RIPK4			
TRIM45				RPRM			
TRIM74				RUNX3			
UBE2DNL				SCRN1			
VAMP5				SERP2			
VWCE				SFRP1			
WDR69				SLC4A4			
ZIK1				SLIT2			
ZNF578				SOHLH2			
ZNF75A				SOX11			
				SPESP1			
				SPINT2			
				SPOCK2			
				SST			
				STAG3			
				STK32B			
				STOM			
				TAC1			
				TCF15			
				TDRD12			
				TDRD9			
				TMEM171			
				TMEM88			
				TNFRSF10C			
				TNFRSF25			
				TRIM45			
				TRIM74			
				TSPYL5			
				TSPYL6			
				VWCE			
				ZNF167			
				ZNF300			
				ZNF331			
				ZNF502			
				ZSCAN18			

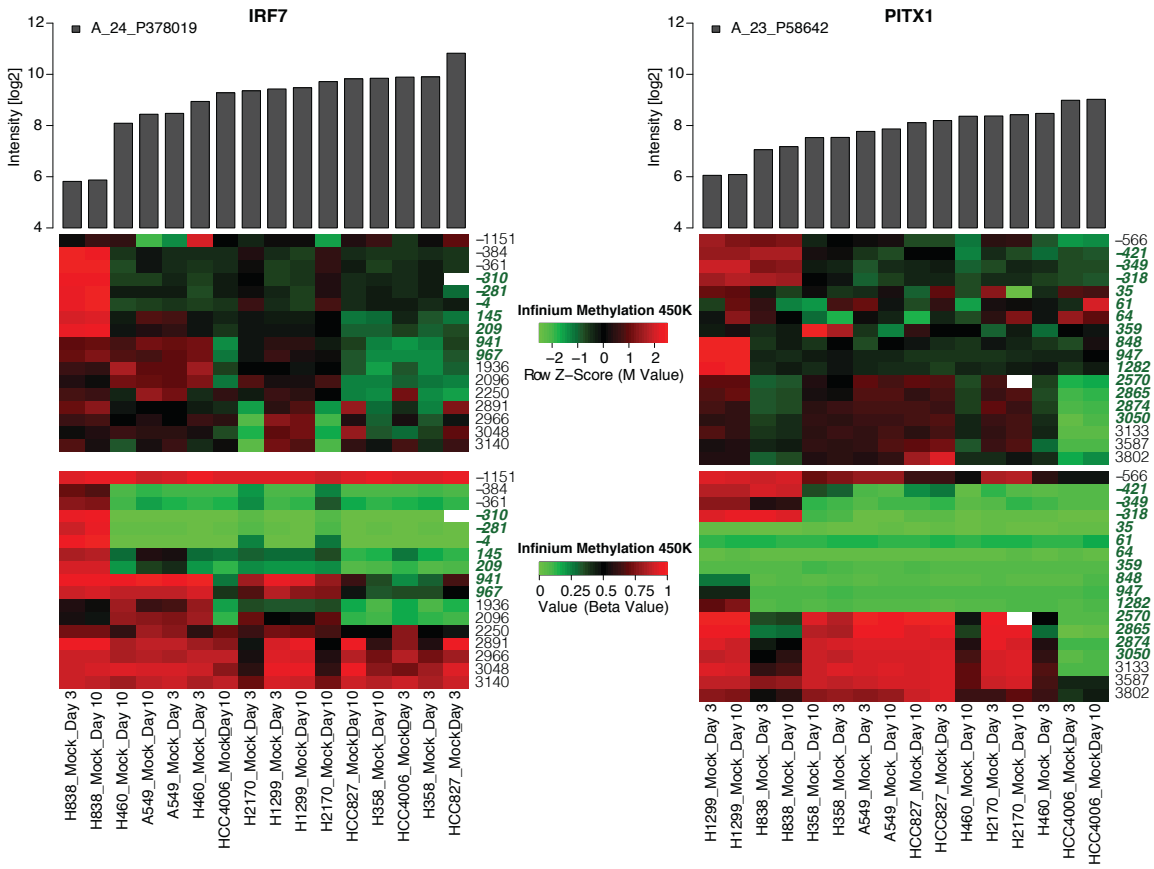


Figure 3.9. Plots (left for *IRF7* and right for *PITX1*) are generated similarly as in Figure 3.3B for mock-treated samples in the eight NSCLC cell lines.

Top panels: bar plots of gene expression levels (log_2 intensity). Bottom panels: heat maps for corresponding DNA methylation levels (red for more and green for less methylated) plotted across the promoter region as in Figure 3.3B with the probe positions from the transcription start sites on the Y-axis with the CpG island probes marked in green. Samples are ordered by their expression levels (from left to right, low to high), as displayed in the top panels.

Table 3.2. PSCAN analysis of the association of the transcription factor IRF7 with gene lists from Figure 3.4A and 3.8B.

PSCAN assigns a measure of statistical association for a transcription factor with a given gene list. The table reports estimated p-values using the TRANSFAC database of co-expressed and co-regulated genes when analyzing -450 to +50 base pairs of promoter sequences in the defined gene lists.

	IRF7
4-Fold Upregulated in H2170	$7.6\ e^{-18}$
Reactome Interferon Alpha Beta	$4.7\ e^{-14}$
Death Genes	$5.7\ e^{-08}$
Mouse IRF7	$2.0\ e^{-05}$
Inflammasome	$6.6\ e^{-05}$
KEGG RIGI Like Receptor Signaling	$1.3\ e^{-04}$
KEGG Toll Like Receptor Signaling	$4.5\ e^{-04}$
Reactome IFN Gamma	$3.3\ e^{-03}$
Immune Tolerance Regulation	0.042
KEGG Antigen Processing and Presentation	0.33
KEGG NFKB	0.40
Stress Response	0.87

Table 3.3. Genes 4-fold or more up-regulated in H2170 at day 10 (7 days post 5-Aza-CR withdrawal).

Gene symbols are indicated in the first column of each panel, the fold changes of the gene expression levels by 5-Aza-CR at day 10 in H2170 are in the second column, availability of RNA-Seq expression data is indicated in column 3, and the chromosomes on which the gens reside are indicated in column 4. Only genes that reside on autosomes and have available TCGA RNA-Seq data are used to plot Figure 3.8B.

Gene Symbol	Expression Fold Change (Aza:Mock)	Genes with Available RNA-seq Data from TCGA	Chr	Gene Symbol	Expression Fold Change (Aza:Mock)	Genes with Available RNA-seq Data from TCGA	Chr	Gene Symbol	Expression Fold Change (Aza:Mock)	Genes with Available RNA-seq Data from TCGA	Chr
ADAM8	13.1	Yes	10	HLA-F	4.1	Yes	6	PARP12	5.5	Yes	7
ANKRD1	8.1	Yes	10	IFFO1	4.0	Yes	12	PARP14	4.2	Yes	3
ANXA8L2	6.0	Yes	10	IFI16	7.8	Yes	1	PARP9	7.3	Yes	3
AREG	4.1	Yes	4	IFI27	22.2	Yes	14	PLAT	4.3	Yes	8
ARHGDIB	5.8	Yes	12	IFI44	25.0	Yes	1	PLK2	4.2	Yes	5
B2M *	3.9	Yes	15	IFI44L	23.3	Yes	1	PLSCR1	8.2	Yes	3
BATF2	4.4	Yes	11	IFI6	5.2	Yes	1	PRSS23	4.7	Yes	11
BST2	7.9	Yes	19	IFIH1	6.8	Yes	2	PSMB9	4.6	Yes	6
C4BPB	10.6	Yes	1	IFIT1	26.7	Yes	10	REC8	4.6	Yes	14
CCL5	5.0	Yes	17	IFIT2	8.4	Yes	10	RPSA	4.5	Yes	3
CLDN1	4.1	Yes	3	IFIT3	21.2	Yes	10	RTP4	8.0	Yes	3
CLDN6	4.4	Yes	16	IFIT5	8.1	Yes	10	S100A16	4.1	Yes	1
CMPK2	4.7	Yes	2	IFITM1	8.8	Yes	11	SAMD9L	8.7	Yes	7
COX7B2	6.7	Yes	4	IFITM2	9.3	Yes	11	SP110	6.1	Yes	2
CST6	11.7	Yes	11	IFITM3	13.4	Yes	11	SPANXA1	26.8		X
CSTA	5.8	Yes	3	IFITM4P	12.7	Yes	6	SPANXB2	50.8	Yes	X
CT45A1	18.9	Yes	X	IGFBP3	6.8	Yes	7	SPANXD	25.3		X
CT45A5	20.7	Yes	X	IRF7	7.9	Yes	11	SPINK1	5.3	Yes	5
CXCL11	5.3	Yes	4	IRF9	10.7	Yes	14	SRGN	4.7	Yes	10
DDX58	13.2	Yes	9	ISG15	12.2	Yes	1	STAT1	8.1	Yes	2
DDX60	24.6	Yes	4	KRT17	6.9	Yes	17	TAP1	5.6	Yes	6
DDX60L	9.5	Yes	4	KYNU	4.3	Yes	2	TDRD12	12.3	Yes	19
DEFB1	5.1	Yes	8	LAMP3	7.1	Yes	3	TFPI2	4.5	Yes	7
DHRS2	6.8	Yes	14	LEPREL1	4.5	Yes	3	TGFBI	4.0	Yes	5
DHX58	7.1	Yes	17	MMP13	6.2	Yes	11	TKTL1	6.9	Yes	X
EPSTI1	5.4	Yes	13	MT1B	7.1	Yes	16	TNFAIP2	8.5	Yes	14
F3	5.1	Yes	1	MT1G	5.8	Yes	16	TPM2	5.3	Yes	9
FHL2	4.3	Yes	2	MT1H	8.6	Yes	16	TRIM22	5.6	Yes	11
FLJ13744	5.4		6	MT1X	9.7	Yes	16	UBD	4.8	Yes	6
FMR1NB	6.8	Yes	X	MT2A	11.7	Yes	16	UCA1	10.2	Yes	19
FN1	9.7	Yes	2	MX1	25.7	Yes	21	USP18	9.0	Yes	22
FSTL1	4.6	Yes	3	MX2	5.2	Yes	21	VCX	20.5	Yes	X
G0S2	8.2	Yes	1	NFE4	9.4			VCX2	12.2	Yes	X
GTSF1	6.4	Yes	12	NLRP2	4.5	Yes	19	VCX3A	19.7	Yes	X
HCLS1	11.2	Yes	3	OAS1	7.5	Yes	12	VCY	8.5	Yes	Y
HERC5	6.1	Yes	4	OAS2	8.4	Yes	12	XAF1	8.3	Yes	17
HERC6	6.5	Yes	4	OAS3	12.9	Yes	12	XK	5.0	Yes	X
HLA-C	4.1	Yes	6	OASL	9.1	Yes	12	ZBED2	4.9	Yes	3

* 3.9 fold upregulated in H2170

We next extrapolated, as was done for the RASSF1 signature earlier, all of our immune-related data to primary NSCLC samples in TCGA. We first approached this by examining those genes found to be increased by 4-fold or more by 5-Aza-CR in the H2170 cell line which harbors the 9-fold increase for *IRF7* (Figure 3.4J). The first important finding is that in both LUSC and LUAD we find a general pattern of concordant lower expression of immune-related genes (Figure 3.8B). Moreover, many of these are likely IRF7 targets identified by promoter sequences in PScan analyses, and/or according to the literature (Table 3.2). Also, for each histology, there are subgroups with especially low expression levels which concordantly cluster to either all the 5-Aza-CR responsive genes in H2170 (Figure 3.8B), just the 5-Aza-CR up-regulated genes from Figure 3.4C-J (Figure 3.10), or to 5-Aza-CR up-regulated genes enriched in immune-related pathways listed in Figure 3.4A (Figure 3.8). Importantly, *IRF7* has lower basal expression, increased transcription start site DNA methylation, and low expression tracking with low immune expression subgroups of LUSC (Figure 3.8B & 3.11). Interestingly, the IRF suppressor gene, *PITX1*, is expressed in most LUSC, but has transcription start site DNA methylation and lower expression in LUAD (Figure 3.8C & 3.11). Finally, and importantly, expression levels of *PD-L1*, the key tumor ligand targeted in the anti-checkpoint immunotherapy trials, tracks quite well with the above immune evasion signature in both LUSC and LUAD, as especially well visualized in heat maps for immune related pathways (Figure 3.11).

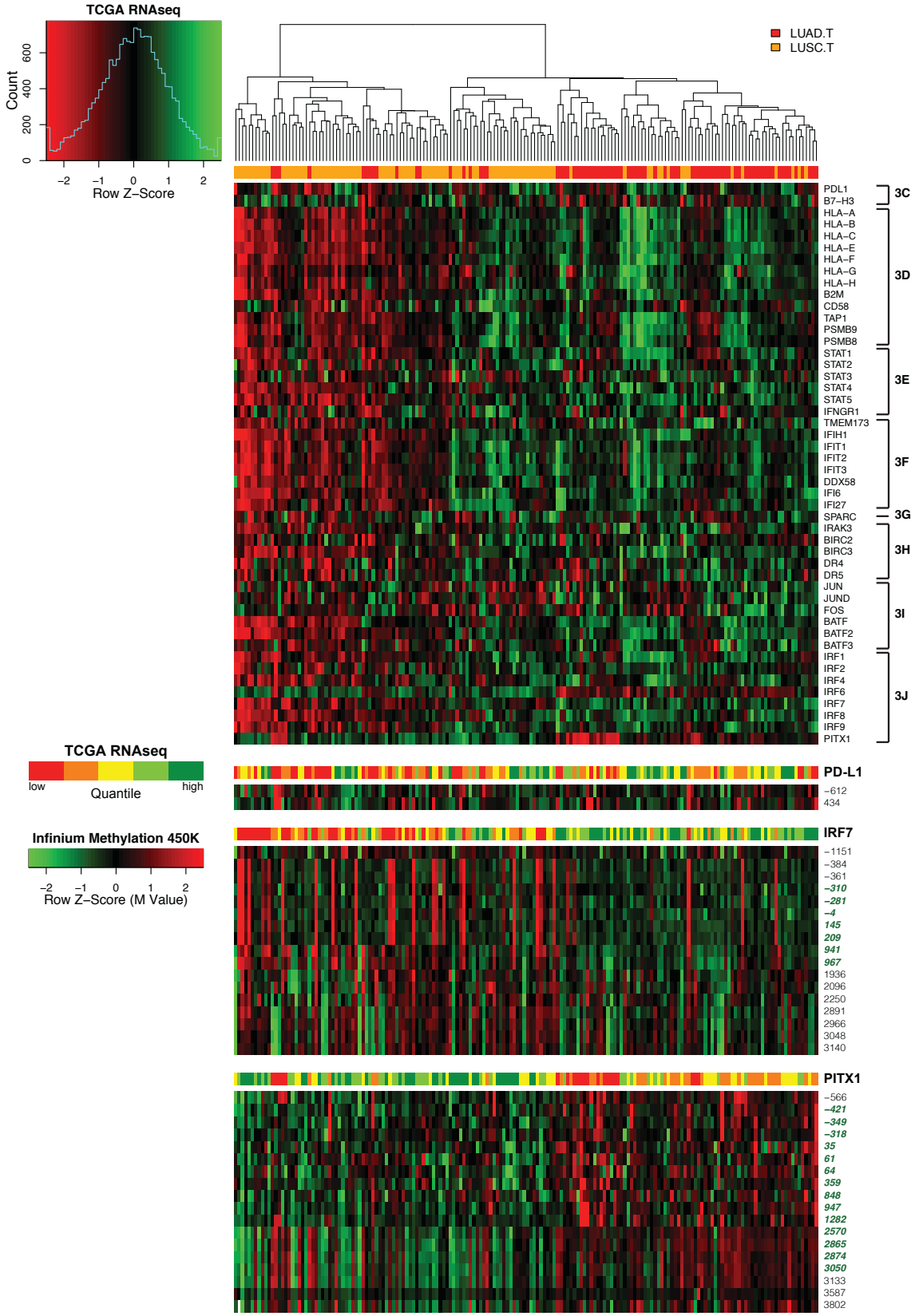

TCGA RNAseq
Count
0
200
400
600
−2
−1
0
1
2
Row Z−Score
LUAD.T
LUSC.T
PDL1
B7−H3
HLA−A
HLA−B
HLA−C
HLA−E
HLA−F
HLA−G
HLA−H
B2M
CD58
TAP1
PSMB9
PSMB8
STAT1
STAT2
STAT3
STAT4
STAT5
IFNGR1
TMEM173
IFIH1
IFIT1
IFIT2
IFIT3
DDX58
IFI6
IFI27
SPARC
IRAK3
BIRC2
BIRC3
DR4
DR5
JUN
JUND
FOS
BATF
BATF2
BATF3
IRF1
IRF2
IRF4
IRF6
IRF7
IRF8
IRF9
PITX1
3C
3D
3E
3F
3G
3H
3I
3J
TCGA RNAseq
low
Quantile
high
PD-L1
−612
434
Infinium Methylation 450K
−2
−1
0
1
2
Row Z−Score (M Value)
IRF7
−1151
−384
−361
−310
−281
−4
145
209
941
967
1936
2096
2250
2891
2966
3048
3140
PITX1
−566
−421
−349
−318
35
61
64
359
848
947
1282
2570
2865
2874
3050
3133
3587
3802

Figure 3.10. TCGA NSCLC samples cluster into subgroups with concordant low expression levels of genes up-regulated by 5-Aza-CR in Figure 3.4C-J.
TCGA heat map for expression of genes (top panel) from Figure 3.4C-J (shown as panels on the Y-axis) that are up-regulated greater than 0.5-fold (log2 scale) in at least one of the eight NSCLC cells lines after treatment with 5-Aza-CR compared to mock at day 10. Panels beneath show expression of *PD-L1*, *IRF7* and *PITX1* with correlating DNA methylation heat map plotted exactly as in Figure 3.4C.

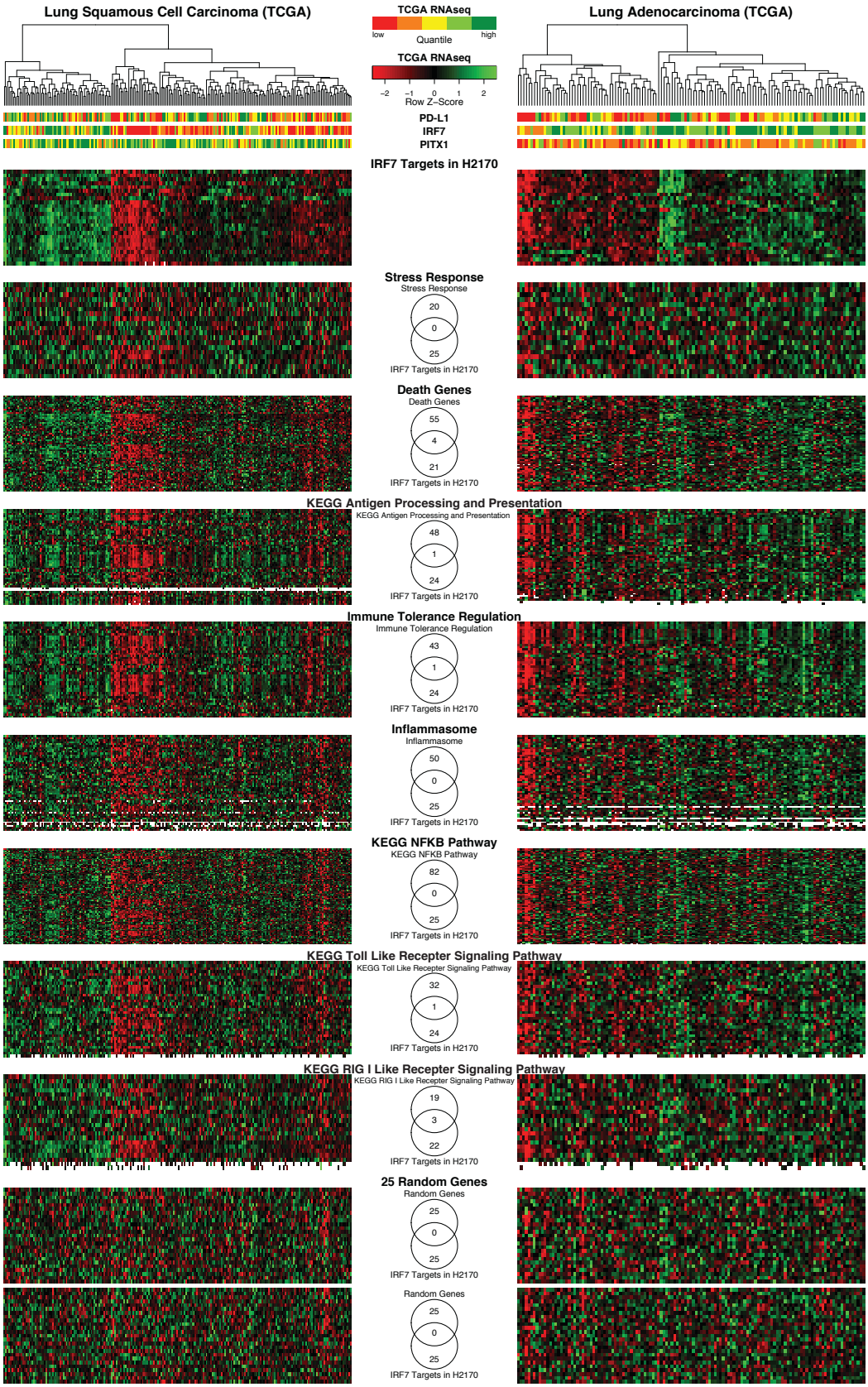

Lung Squamous Cell Carcinoma (TCGA)
TCGA RNAseq
low
high
Quantile
TCGA RNAseq
−2
−1
0
1
2
Row Z−Score
Lung Adenocarcinoma (TCGA)
PD-L1
IRF7
PITX1
IRF7 Targets in H2170
Stress Response
Stress Response
20
0
25
IRF7 Targets in H2170
Death Genes
Death Genes
55
4
21
IRF7 Targets in H2170
KEGG Antigen Processing and Presentation
KEGG Antigen Processing and Presentation
48
1
24
IRF7 Targets in H2170
Immune Tolerance Regulation
Immune Tolerance Regulation
43
1
24
IRF7 Targets in H2170
Inflammasome
Inflammasome
50
0
25
IRF7 Targets in H2170
KEGG NFKB Pathway
KEGG NFKB Pathway
82
0
25
IRF7 Targets in H2170
KEGG Toll Like Recepter Signaling Pathway
KEGG Toll Like Recepter Signaling Pathway
32
1
24
IRF7 Targets in H2170
KEGG RIG I Like Recepter Signaling Pathway
KEGG RIG I Like Recepter Signaling Pathway
19
3
22
IRF7 Targets in H2170
25 Random Genes
Random Genes
25
0
25
IRF7 Targets in H2170
Random Genes
25
0
25
IRF7 Targets in H2170

Figure 3.11. Immune pathways cluster with subgroups of NSCLC defined by expression of IRF7 transcription factor targets.
Left panels are LUSC and right panels are LUAD primary samples. *PD-L1*, *IRF7*, and *PITX1* expression are depicted on the top bar panels by quantile expression. Sample ordering is established by unsupervised clustering of TCGA samples based on expression levels of IRF7 targets derived from H2170. Sample order is maintained across all other bar panels and heat maps demonstrating that subgroups defined by IRF7 transcription factor targets cluster tightly with multiple immune pathways. Venn diagrams depict the overlap with the IRF7 targets up-regulated in H2170 of the most enriched genes in each pathway. A complete table showing the overlaps is available in Table 3.4. That the observed clustering pattern is not due to chance is illustrated by using a random set of 25 genes each in the bottom two panels. All color codes for expression are as per Figure 3.8B-C.

Table 3.4. Overlaps of gene lists utilized to plot the heat maps in Figure 3.11.

	Stress Response	Death Genes	KEGG Antigen Processing and Presentation	Immune Tolerance Regulation	Inflammasome	IRF7 Targets in H2170	KEGG NFKB Pathway	KEGG Toll Like Recepter Signaling Pathway	KEGG RIG I Like Recepter Signaling Pathway
Stress Response	20	-	-	-	-	-	-	-	-
Death Genes	1	59	-	-	-	-	-	-	-
KEGG Antigen Processing and Presentation	0	1	49	-	-	-	-	-	-
Immune Tolerance Regulation	1	2	13	44	-	-	-	-	-
Inflammasome	0	14	0	2	50	-	-	-	-
IRF7 Targets in H2170	0	4	1	1	0	25	-	-	-
KEGG NFKB Pathway	0	9	0	2	6	0	82	-	-
KEGG Toll Like Recepter Signaling Pathway	2	3	0	3	5	1	9	33	-
KEGG RIG I Like Recepter Signaling Pathway	0	6	1	0	1	3	6	6	22

All of these above data are potentially highly relevant to recent occurrences for our recently completed clinical trial. Five patients who received the epigenetic therapy, after progressing, were placed on trials for immunotherapy targeting the immune tolerance checkpoint (Juergens et al., 2011). Each patient received medical benefit which can be placed into context for this immunotherapy approach alone. In this regard, 85 patients with the same stages of LUAD and LUSC as in our epigenetic therapy trial received monotherapy with anti-PD-1 or anti-PD-L1 and durable objective response rates (ORRs) occurred for 10% of patients for anti-PD-L1 and 18% for anti-PD-1. An additional 12% and 7% of patients demonstrated stable disease (SD) of $\geq$ 24 weeks, respectively (Brahmer et al., 2012; Topalian et al., 2012). While these responses are the highest ever reported for immunotherapy in advanced lung cancer (Brahmer et al., 2012; Topalian et al., 2012), the five patients receiving this immunotherapy after epigenetic therapy (patient data in Table 3.5), treated either with anti-PD-1 or anti-PD-L1have all shown treatment benefit lasting at least 24 weeks, and three of five have experienced major radiologic responses to subsequent immunotherapy, all of which are ongoing from 18 to 24+ months (Figure 3.12). Obviously, the true potential of combining epigenetic therapy with cytotoxic or immunotherapies warrants follow-up trials, which are soon to start. However, these trials could benefit greatly from the laboratory data contained in our present study as it provides potential insights into mechanisms that might underlie the above promise and a path forward to generate biomarker strategies which can be tested in the upcoming validation trial.

Table 3.5. Patient data for 5 patients treated with combination epigenetic therapy consisting of 5-Aza-CR and Entinostat prior to single agent anti-PD-1 or anti- PD-L1 directed immune-checkpoint blockade.

Patient	1	2	3	4	5
Gender	Male	Male	Female	Female	Male
Age	62	62	64	68	62
Histology (mutation, if known)	LUAD	LUSC	LUAD (KRAS)	LUSC	LUSC
Number of Treatment before Epigenetic Treatment	3	1	1	1	2
Best Response to Any Prior Treatment *	SD	PR	SD	PD	PR
Duration of Treatment with 5-Aza-CR & Entinostat prior to Immune Treatment (months)	2	2	6	2	8
Best Response to Epigenetic Treatment *	PD	PD	MR	PD	PR
Type of Immune Treatment	Anti-PD1	Anti-PD1	Anti-PD-L1	Anti-PD1	Anti-PD-L1
Duration of Immune Treatment	19+ months (ongoing)	22.5+ months (ongoing)	10.5+ months (ongoing)	8.25 months	8.5 months
Response to Immune Therapy	PR (-57%)	PR (-73%)	PR (-64%)	SD (+11%)	SD (+18%)

* SD = Stable Disease
PR = Partial Response
PD = Progressive Disease
MR = Mixed Response (some lesion responding, some progressing)

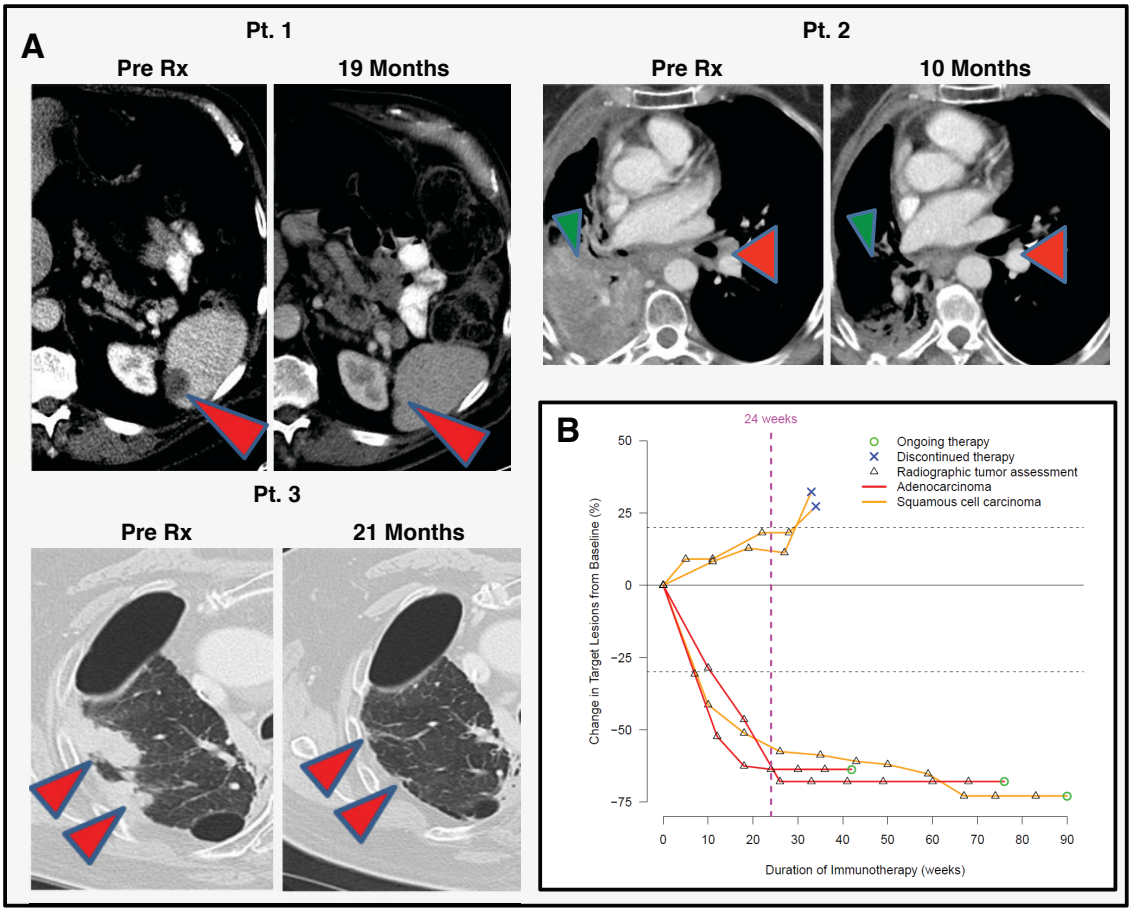


Figure 3.12. Outcomes for five NSCLC patients treated with anti-PD1 immune checkpoint inhibitor immunotherapy after epigenetic therapy.
(A) CT scans for 3 patients who responded to anti-PD1 immune checkpoint therapy. Tumor lesions are highlighted by red arrows (Pt.1 with splenic metastasis resolving by 19 months on immunotherapy; Pt.2 with involved lymph node in left chest and green arrow denoting an area of the right lung collapsed behind airway obstruction due to tumor both resolving over a 10-month period on immunotherapy; Pt.3 lung parenchymal tumor resolving over 21 months on immunotherapy) (B) Spider plot of RECIST measurements from serial CT scans over time for 5 patients who received anti-PD1 immune checkpoint therapy after epigenetic therapy. A decrease in tumor measurements of 30% or more qualifies as RECIST criteria response (green circles). An increase in tumor measurements of 20% or more qualifies as disease progression (blue X's). 24 weeks denoted by the dashed vertical line represents the time at which a patients with less than 20% increase in tumor measurements are considered stable disease as defined by the published trials of anti-PD1 therapy (Brahmer et al., 2012; Pardoll, 2012; Topalian et al., 2012).

We must address, lastly, a key issue, concerning mechanisms that may account for all of the pre-clinical data we have presented. These are probably quite complex, dose-related actions of 5-Aza-CR. As we show, the drug induces a cascade of direct and indirect effects within pathways. However, this drug is best known for inducing DNA demethylation through direct inhibition of the catalytic sites and triggering degradation of biologically active DNMTs (Gabbara and Bhagwat, 1995; Santi et al., 1984). We, thus, compared findings from NSCLC cell lines to HCT116 colon cancer cells and HCT116 double knock out (DKO) cells that have been genetically disrupted to give severe haplo-insufficiency of DNMT1, and complete absence of DNMT3B, which are enzymes for DNA methylation maintenance and *de novo* DNA methylation, respectively (Rhee et al., 2002). Gene expression alterations in DKO versus wild type HCT116 (Figure 3.13A-H) are remarkably similar to the 5-Aza-CR induced changes (Figure 3.4C-J) in NSCLC cells. Thus, off target effects of high dose 5-Aza-CR, including incorporation into RNA and DNA as an abnormal nucleotide (Stresemann and Lyko, 2008), do not appear to be required for many of the drug's effect that we have defined.

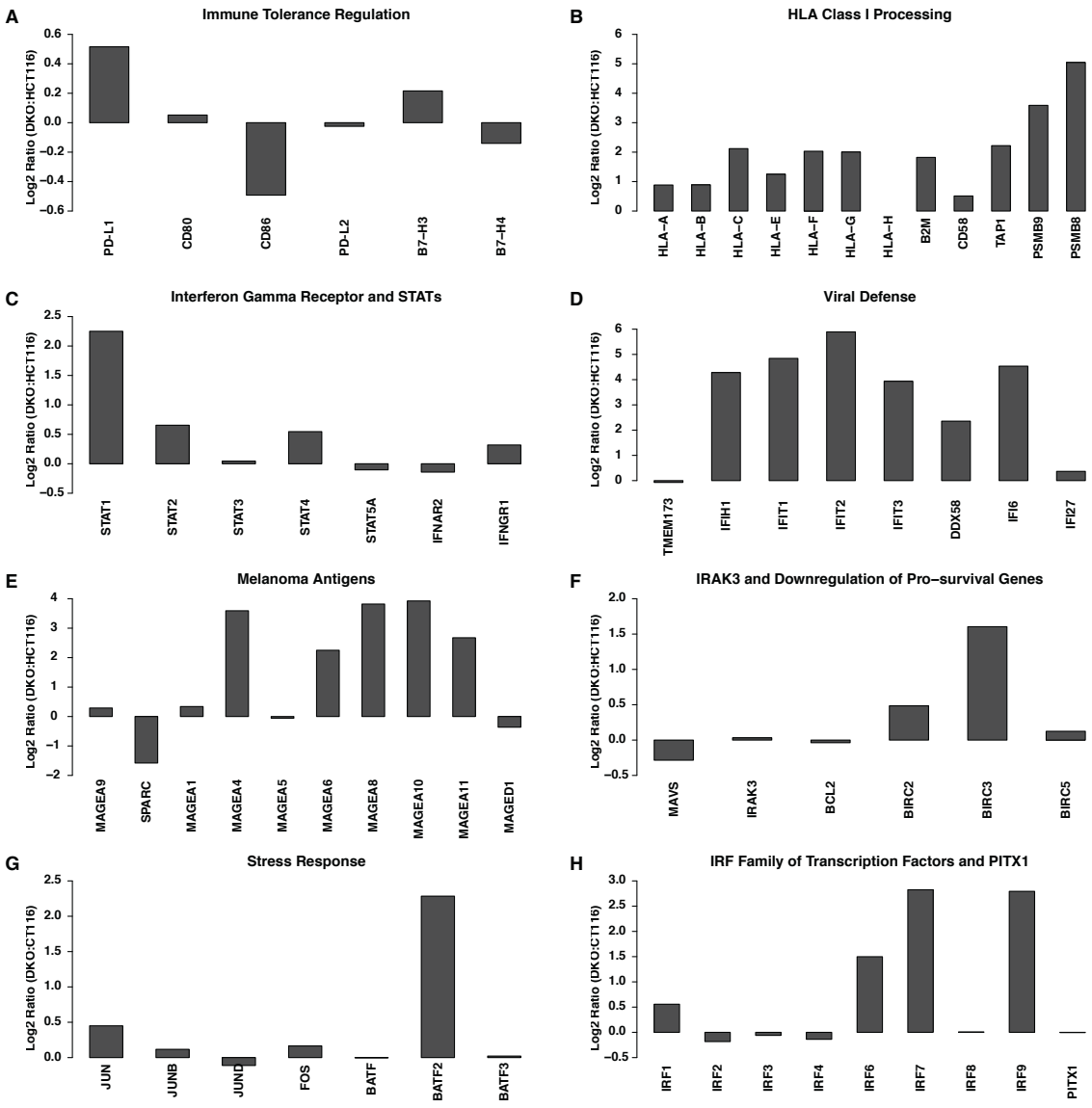


Figure 3.13. Gene expression alterations in DKO versus wild type HCT116.
The gene expression differences are given as the $\log_2$ ratio of expression in DKO over wild type cells (Y-axis) and the gene panels, A-H correspond to panels C-J in Figure 3.4 for the NSCLC cell lines treated with 5-Aza-CR.

3.5 Discussion

In the present work, we have taken a novel track for deriving pre-clinical understanding of tumor cell responses that are potentially highly relevant to emerging promise of a drug in the clinic. By matching relevant drug dose in cell lines to not only tumor responses *in vivo* but also to genome-wide DNA methylation and gene expression responses, we see robust pathway responses important for the pre-clinical agent effects. Then, by matching these profiles to the extensive genomic data in the TCGA project, we observe what could prove to be key signatures relevant to primary cancers.

With respect to the specific drug studied, we derive potentially valuable insight into how an agent targeting epigenetic abnormalities in NSCLC may have clinical efficacy. Our results are relevant for existing clinical trial data indicating that a regimen employing 5-Aza-CR may yield patient responses to the therapy alone and/or re-sensitization to standard chemotherapies and new forms of immunotherapy. Moreover, we observe the possibilities that the pathways and genes which emerge from our data may interact in contributing to all of these clinical responses. As noted previously, the clinical responses discussed obviously need to be validated in subsequent clinical trials which are about to begin. Thus, we will be testing patients with advanced NSCLC of both LUAD and LUSC histologies, for continued responses to epigenetic therapy alone and formally to see if we extend the promise for sensitization to standard chemotherapy and to immune tolerance checkpoint therapy.

A key issue to mention is that in our clinical trials we have employed the HDAC inhibitor Entinostat in combination with 5-Aza-CR (Juergens et al., 2011). This agent

could of course be playing a major role in any patient responses of the types under discussion. This drug is known to up-regulate many of the genes we have studied, and may be especially involved in host immune responses (Claus et al., 2005; Fonsatti et al., 2007; Simova et al., 2011). These possibilities must be pursued further in subsequent studies, and especially, host immune cell responses will be closely monitored in our subsequent clinical trial of anti-PD1 therapy. The HCT116-DKO model has allowed us to look at a comparison between a genetic knock out of DNMT's, the 5-Aza-CR induced responses, and responses to the potent HDACi trichostatin (TSA). For essentially all of the genes we have examined, the genetic disruption of DNMT's and 5-Aza-CR treatment are the maneuvers which either singularly or most potently induce gene up-regulation in the HCT116 tumor cells (data not shown). We now need to extend these data in future studies and establish the role of both drugs for host immune responses in addition to those within the epithelial tumor cell.

In summary, the data we present in the present work now give us a targeted approach to query, in pre- and post- tumor biopsies, whether events like the status of *RASSF1* and *IRF7* DNA methylation, demethylation, and gene expression changes, plus expression changes in all the cell cycle, stem cell, apoptosis, and immune regulation pathways we have observed track with any patients benefits which may continue to emerge. If such patient efficacy is seen we may be developing a crisp biomarker approach for determining which individuals are likely to benefit from the epigenetic therapy approaches we are testing in clinical trials. Finally, the gene and pathway responses we are observing have important implications for future clinical trial design and provide

strong clues as to how to employ drugs like 5-Aza-CR in even more ways to sensitize to chemotherapy and targeted agents.

3.6 Materials and Methods

3.6.1 Clinical Data

Institutional review board approved informed consent signed by each patient allowed the collection of clinical data following treatment on trial with epigenetic therapy. Relevant data were obtained by chart review. Representative images demonstrating responses to therapy were obtained from computed tomography series employed in the assessment of patient responses to anti-PD1 or anti-PD-L1 directed immune-checkpoint therapy. Assessment of response to treatment was performed by a single reference radiologist who employed (RECIST 1.0) to generate measurements for target lesions to be followed over the course of therapy. Change in target lesions from baseline (%) is calculated by summing the diameter of all target lesions at each radiographic tumor evaluation and calculating percentage change at a given time point ([(Target Lesion SumTimepoint X/ Target Lesion SumBaseline)-1]*100)

Though patients were not enrolled in a single, randomized clinical trial to test the concept of whether priming with epigenetic therapy augments treatment efficacy of PD-1 directed immune checkpoint inhibition the outcomes for the 5 patients was statistically different from patients receiving immunotherapy alone ($p < 0.02$, Fisher's Exact Test) and should be taken as preliminary evidence supporting further scientific and clinical investigation.

3.6.2 Solid Tumor Xenograft Tumorigenicity Assay

The measurements of tumor xenografts was performed exactly as per our previous publication (Tsai et al., 2012). Specifically, for this study, NSCLC lines H838, H1299, H358, H2170, H358, H460, HCC827, and HCC4006 cells were pretreated with 500 nM azacytidine or PBS (Mock) for 72 hr followed by another 7 days in culture without drug. Harvested cells were injected (1×10^6) subcutaneously into a single flank of five 4- to 6-week-old male NOD/SCID mice. Serial tumor measurements were obtained on a weekly basis after tumor xenograft injection.

3.6.3 TCGA Samples

Level 3 RNA-Seq data (Illumina HiSeq RNA-Seq platform, Illumina, Inc., San Diego, CA, USA) were downloaded for 353 NSCLC samples (129 LUAD / 224 LUSC) and 54 adjacent non-tumor lung tissue samples from the TCGA Data Portal (https://tcga-data.nci.nih.gov/tcga/). Similarly, level 1 DNA methylation data (Illumina Infinium HumanMethylation450 BeadChip, Illumina, Inc., San Diego, CA, USA) were downloaded for 353 NSCLC samples (222 LUAD / 149 LUSC) and 74 adjacent non-tumor lung tissue samples. Among these, data for 174 NSCLC samples (80 LUAD / 94 LUSC) and 21 adjacent non-tumor lung tissue samples were available on both of the above platforms.

3.6.4 RNA-Seq Data Analysis

We used TCGA level 3 RNA-Seq data that were already normalized and quantified at gene levels, and were presented as RPKM values (Reads Per Kilobase per Million mapped reads). To construct heat maps: 1) Values of 0 (indicating no reads observed for

a gene) in the RPKM data were set to NA; 2) the remaining RPKM values were log 2 transformed; 3) genes from X and Y chromosomes were removed; and 4) heat maps were made using the "heatmap.2" function in "gplots" package from CRAN (Warnes, 2012), being centered and scaled in the row direction, and using the default functions for computing distance and hierarchical clustering (or being specifically ordered in column according to the order of other heat maps). Expression spectrums for individual genes were displayed in five quantile intervals following the order of associated heat maps of the RNA-Seq data.

3.6.5 Infinium DNA Methylation Data Analysis

TCGA level 1 DNA methylation data contained the raw binary intensity data files. Raw data files were imported into R (http://www.r-project.org) to calculate beta values (beta value $_{\text{Infinium}}$ = M / [U + M], M: mean intensities of the Methylated bead types, U: mean intensities of the Unmethylated bead types), M values (M value $_{\text{Infinium}}$ = $\log_2$ [M / U]) and detection p-values (calculated by comparing probes to negative control probes to determine if signals are significantly different from the background) using the "methylumi" package from Bioconductor (Davis et al., 2012). beta values and M values for probes with detection p-value > 0.05 were considered not significantly different from background and were masked as NA.

TCGA methylation data were first assessed for batch effects by principle component analysis (PCA) on the M values. To accomplish this, data points from X chromosome and Y chromosome as well as data points that are associated with SNPs (Single Nucleotide Polymorphisms) were removed, and the first two principle components are used for

plotting (Figure 3.14). To better illustrate the relationship between different batches, plots were enhanced by displaying the centroid of each batch (Figure 3.14C & D), where centroid was calculated by taking the mean of all the samples in the corresponding batch. From the analysis we can conclude that there is distinct difference in TCGA methylation data between NSCLC samples and adjacent non-tumor lung tissue samples, as well as between lung adenocarcinoma (LUAD) and squamous cell carcinoma (LUSC) samples, and there were no significant batch effects throughout the studies (Figure 3.14).

Spearman's correlation coefficients between methylation (beta value of probe, Illumina Infinium HumanMethylation450 BeadChip) and gene expression (RPKM value of gene, Illumina HiSeq RNA-Seq platform) were calculated using TCGA samples with available data on both platforms. For a particular gene, only methylation probes that have a negative Spearman's correlation coefficient and a adjusted p-value (FDR) for the coefficient < 0.01 were considered informative and their relative distances to the corresponding transcriptional start site (TSS) of the genes were calculated from genomic coordinates obtained from the UCSC genome browser (http://genome.ucsc.edu). Heat maps of the M values of informative probes were made using the "heatmap.2" function in "gplots" package from CRAN (Warnes, 2012), being centered and scaled in the row direction, and ordered according to the associated heat maps of the RNA-Seq data in column and to the relative distances to TSS in row.

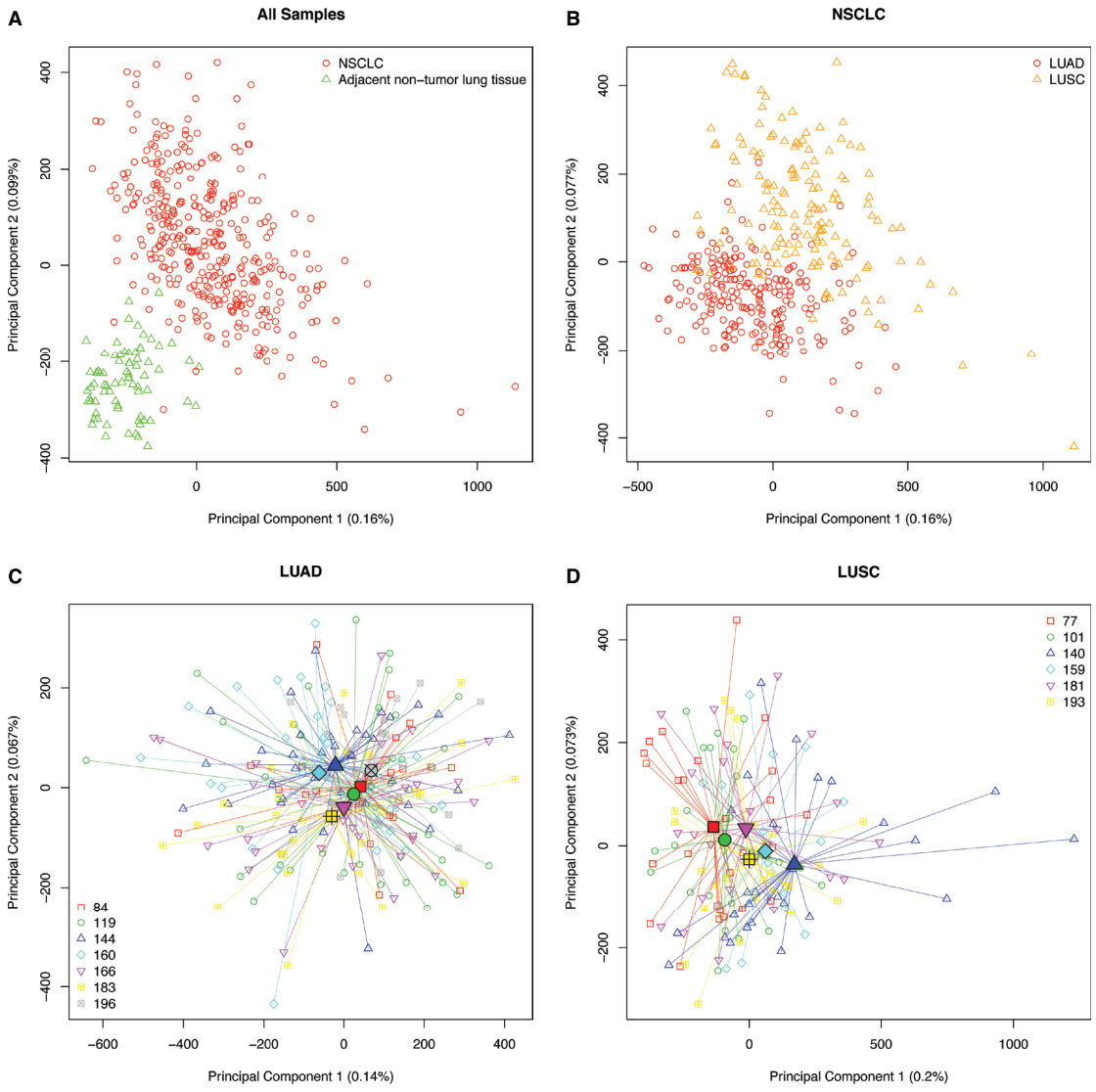


Figure 3.14. PCA analysis of TCGA level 1 DNA methylation data.
(A) Analysis of all samples indicates a distinct separation of NSCLC samples (red) from the adjacent non-tumor tissue samples (green). (B) When the analysis is performed only on the NSCLC samples, a distinct separation was observed between LUAD (red) and LUSC (orange). PCA analysis of only LUAD samples (C) and LUSC (D) indicates limited batch effects, which are obvious by the close distribution pattern of the centroids (larger solid symbols), with batches indicated by different colors.

For *in vitro* DNA methylation values, DNA was extracted from cell lines that were either untreated or treated with 5-Aza-CR at day 3, at the end of treatment, and day 10 (7 days post end of treatment) and analyzed by the Illumina Infinium HumanMethylation450 BeadChips (Illumina, Inc., San Diego, CA, USA). Raw data were imported into R using the "methylumi" package from Bioconductor (Davis et al., 2012). Data points for probes with detection p-value > 0.05 were masked as NA. Δ beta values (Δbeta value=beta value$_{Aza}$-beta value$_{Mock}$) were calculated and used to make boxplots. Heat maps were made similarly like those for the TCGA data using informative probes defined by the TCGA data.

3.6.6 Expression Microarray Data

For in-vitro RNA extracted from cell lines treated with 5-Aza-CR, analyses were done at exactly the same time points as for DNA methylation above. Analyses from wild type colon cancer, HCT116 cells, and genetic knockout counterparts for DNA methyltransferases (DKO cells) were also performed. Expression microarrays were carried out using Agilent Human 4× 44K expression arrays (Agilent Technologies, Santa Clara, CA, USA, Cat#: G4112F). Within-array and between - array normalization was performed using Loess and Aquantile normalization, respectively (Smyth and Speed, 2003). Median of the M values (M value $_{Expression}$ = $\log_2$ [5-Aza-CR / Mock] OR $\log_2$ [DKO / HCT116]) was determined for multiple probes associated with the same gene.

3.6.7 Gene Set Enrichment Analysis (GSEA)

For each of the eight lung cancer cell lines (H838, H1299, H358, H1270, A549, H460, HCC4006, HCC827) a ranked gene list was created (genes were sorted by decreasing M value). These eight ranked gene lists were entered in the GSEA tool (Subramanian et al., 2005) and the enrichment of both Kegg (Goto et al., 1997) and Reactome (Joshi-Tope et al., 2005) pathways in these lists was calculated (default parameters). A gene set was selected when it was enriched in any of the eight cell lines (p value < 0.05 and false discovery rate < 0.25). The normalized enrichment scores (NES) for the gene sets in each cell line were used to create the heat maps. When a certain gene set was not significant in a cell line, it was assigned a NES of 0.

3.6.8 Transcriptional Factor Analysis

Expression and methylation data were analyzed to find genes whose re-expression was linked to demethylation after 5-Aza-CR treatment. Genes were selected based on a set of cut-offs, both for the methylation and expression values: A gene was considered to be re-expressed when at day 3 or day 10 the median M value of all the probes linked to that gene was higher than 0.5. Infinium probes were analyzed separately at their distances from the transcription start site for each gene examined. For a probe to be called demethylated, it had to have a beta value higher than 0.5 in the mock treatment and the difference in beta value between mock and 5-Aza-CR treatment had to be higher than 0.25. Only probes that were associated with a CpG island and that were located within 1000 bp upstream and 1000 bp downstream of the transcription start site were used in the analyses. The probes that passed these filters were validated using the TCGA methylation and expression data (see the definition of informative probes in the “Infinium DNA

Methylation Data" section of Methods). Only genes that had an expression-methylation correlation value < -0.25 and a false discovery rate < 0.05 were retained.

To better understand the biological implications of the re-expressed genes, the gene lists were searched for transcription factors. Two human transcription factor lists obtained from Ravasi et al. (Ravasi et al., 2010) and Vaquerizas et al. (Vaquerizas et al., 2009) were combined and the resulting list was matched to the lists of demethylated and re-expressed genes. The targets of *IRF7* from the list of genes that are 4-fold or more up-regulated in H2170 by 5-Aza-CR were similarly identified using the TranscriptomeBrowser database (Ravasi et al., 2010; Vaquerizas et al., 2009).

3.6.9 Flow Cytometry Methods (FACS)

Frozen cells were thawed in 37 degrees celcius and washed once with flow-washing buffer. Aliquots of single-cell suspension were then stained with fluorescent-labeled antibodies for 15 mins at room temperature. Each sample was washed twice and resuspended in flow-washing buffer and analyzed by FACSCalibur. The following antibodies were used: CD274 (12-5983-42 Ebiosciences), HLA abc (12-9983-42 Ebiosciences), CD276 (331606 Biolegend), CD119 (558934 BD), B2 microblogumin (551337BD), CD58 (555921BD).

3.6.10 PSCAN

PSCAN (http://159.149.160.51/pscan/) (Zambelli et al., 2009) is an online software tool that predicts the association of user-defined gene-lists with transcription factors by scanning promoter sequences of co-regulated or co-expressed genes looking for over- or

under-represented motifs. RefSeq IDs of the gene lists were obtained from BioMart (http://www.biomart.org/) and analyzed in PSCAN. Scanned promoter region was -450 to +50 base pairs around the transcription start site and employing TRANSFAC as the database for co-regulated or co-expressed genes.

3.7 Acknowledgements

This paper was supported by grants from CA058184, National Cancer Institute (NCI) and the Stand Up To Cancer (SU2C) and the Samuel Waxman Cancer Research Foundation. We gratefully acknowledge the TCGA consortium for creation of the public database from which we queried RNA-Seq gene expression data and DNA methylation status for selected genes. Drs. Laird (Principal Investigator - USC) and Baylin (co-Principal Investigator - JHU) lead the epigenetic analyses in TCGA, while Dr. Weisenberger leads the efforts to perform the DNA methylation analyses. We thank Kathy Bender for manuscript preparation.

Chapter 4

Conclusions and future directions

The majority of cancers are age-related, caused and/or facilitated by numerous accumulating genetic and epigenetic alterations in response to stimuli experienced during aging, such as tobacco exposure, environment pollution, stress and chronic inflammation. As society advances and the people's average lifespan greatly improves, cancer has become one of the major causes of death in humans. It is of great urgency to understand the underlying mechanisms of cancer initiation and progression, in order to develop more effective strategies to improve cancer prevention, early detection and treatment.

In recent years, epigenetics has been given more and more attention in the cancer research field due to its potentially extensive role in all aspects of cancer and its appealing nature of being relatively easy to reverse by pharmacological agents. Despite the enormous and growing number of papers published every year, we have only scraped the tip of the iceberg in our full understanding of the epigenetic changes in cancer and their causes. Fortunately, with the development of novel technologies and advancement in new analysis methods, especially the widespread use of various genomic-based and bioinformatic approaches, we are able to take cancer epigenetic research to a new level. In this thesis, our studies have taken advantage of these advances to provide further insight into the origins of epigenetic abnormalities in human cancer, and to apply this knowledge to the translational goal of bringing epigenetic therapy to the biggest cancer killer, advanced lung cancer. The conclusions we have reached can be summarized as follows.

In chapter 2, we used a genome-wide chromatin immunoprecipitation approach together with studies of gene expression and DNA methylation to investigate whether the cell stress of rapidly increasing reactive oxygen species, and resultant oxidative damage

to DNA and/or chromatin, induces molecular mechanisms that alter the epigenome. These studies identified that the above stress can, within 30 minutes, trigger oxidative damage-induced increase in size and abundance of, and tightening to chromatin of key constituents of, a pre-existent, transcription repressing, complex containing DNA methyltransferases and members of the cancer specific polycomb repressive complex, PRC4. This complex moves from GC poor genomic regions to damage sites within GC rich ones in gene promoter regions. These findings provide connections between oxidative damage, transient silencing of genes perhaps as a protective mechanism during repair of DNA damage, and reveal how key genes may be vulnerable during these processes to evolve abnormal DNA methylation. The findings, plus work in a mouse model of colon inflammation, suggest how the leading cancer-risk state, chronic inflammation, may induce key epigenetic changes which may play a key role in tumor initiation and facilitate malignant transformation.

Extended studies will be required to further understand the nature of this oxidative damage-induced silencing complex. We must learn more about how it is assembled and recruited to chromatin and why it preferentially targets transcriptionally active and GC rich genomic regions after H_2O2 exposure. Some evidence indicates that targeting of the complex is associated with the distribution of DNA damage sites, thus linking its formation and positioning to chromatin accessibility, types of damage, and the fact that guanine is the most easily oxidized of the four deoxyribonucleosides (Steenken, 1997). Furthermore, an *in vitro* chronic oxidative damage model is needed to examine how the observations we have defined occur over time and how we can extrapolate our studies to *in vivo* events for chronic inflammation models.

In chapter 3, we used a genomic-based approach, and bioinformatic methods, to explore what molecular mechanisms might underlie early promising responses of patients with advanced, non-small cell lung cancer (NSCLC) to therapies employing the DNA demethylating agent, 5-aza-cytidine (5-Aza-CR). By profiling genome-wide gene expression and DNA methylation responses to 5-Aza-CR treatment for eight NSCLC cell lines, we observed in most cell lines a complex multi-faceted up-regulation, including hundreds of genes, of immune-related pathways. Moreover, using the obtained molecular signature, we have been able to specifically query hundreds of primary NSCLC samples in the Cancer Genome Atlas project for how basal expression of these genes groups these cancers. Our findings suggest, especially in the setting of the above patient responses, that epigenetic therapy may sensitize such individuals to immune checkpoint blockade at least, in part, by shifting the tumor cells away from an immune evasion state towards enhanced immune recognition.

A critical aspect of the work, in addition to providing for understanding of the molecular mechanisms underlying the above sensitization, is the potential of the findings to develop biomarker strategies to predict which patients with NSCLC may benefit the most from the therapeutic approach. A larger clinical trial is now just underway in which these strategies can now be tested.

References

Agathanggelou, A., Honorio, S., and Macartney, D. (2001). Methylation associated inactivation of RASSF1A from region 3p21.3 in lung, breast, and ovarian tumors. Oncogene *20*, 1509 - 1518.

Allis, C. D., Jenuwein, T., and Reinberg, D. (2007). Epigenetics, (Cold Spring Harbor, N.Y.: Cold Spring Harbor Laboratory Press).

Altmeyer, M., and Lukas, J. (2013). Guarding against collateral damage during chromatin transactions. Cell *153*, 1431-1434.

Amin, K. S., and Banerjee, P. P. (2012). The cellular functions of RASSF1A and its inactivation in prostate cancer. J Carcinog *11*, 3.

Antequera, F., and Bird, A. (1993). Number of CpG islands and genes in human and mouse. Proc Natl Acad Sci U S A *90*, 11995-11999.

Bannister, A. J., and Kouzarides, T. (2011). Regulation of chromatin by histone modifications. Cell research *21*, 381-395.

Bao, Y., and Shen, X. (2007). INO80 subfamily of chromatin remodeling complexes. Mutation research *618*, 18-29.

Bassett, A., Cooper, S., Wu, C., and Travers, A. (2009). The folding and unfolding of eukaryotic chromatin. Current opinion in genetics & development *19*, 159-165.

Baylin, S., and Bestor, T. H. (2002). Altered methylation patterns in cancer cell genomes: cause or consequence? Cancer Cell *1*, 299-305.

Baylin, S. B. (2011). Resistance, epigenetics and the cancer ecosystem. Nat Med *17*, 288-289.

Baylin, S. B., and Herman, J. G. (2000). DNA hypermethylation in tumorigenesis: epigenetics joins genetics. Trends in genetics : TIG *16*, 168-174.

Bell, A. C., and Felsenfeld, G. (2000). Methylation of a CTCF-dependent boundary controls imprinted expression of the Igf2 gene. Nature *405*, 482-485.

Benedict, C. A., Norris, P. S., and Ware, C. F. (2002). To kill or be killed: viral evasion of apoptosis. Nat Immunol *3*, 1013-1018.

Berger, S. L. (2002). Histone modifications in transcriptional regulation. Current opinion in genetics & development *12*, 142-148.

Berger, S. L., Kouzarides, T., Shiekhattar, R., and Shilatifard, A. (2009). An operational definition of epigenetics. Genes & development *23*, 781-783.

Bernstein, B. E., Mikkelsen, T. S., Xie, X., Kamal, M., Huebert, D. J., Cuff, J., Fry, B., Meissner, A., Wernig, M., Plath, K.*, et al.* (2006). A bivalent chromatin structure marks key developmental genes in embryonic stem cells. Cell *125*, 315-326.

Berry, W. L., and Janknecht, R. (2013). KDM4/JMJD2 histone demethylases: epigenetic regulators in cancer cells. Cancer Res *73*, 2936-2942.

Bhutani, N., Burns, D. M., and Blau, H. M. (2011). DNA demethylation dynamics. Cell *146*, 866-872.

Bidwell, B. N., Slaney, C. Y., Withana, N. P., Forster, S., Cao, Y., Loi, S., Andrews, D., Mikeska, T., Mangan, N. E., Samarajiwa, S. A.*, et al.* (2012). Silencing of Irf7 pathways in breast cancer cells promotes bone metastasis through immune escape. Nat Med *8*, 1224-31.

Bird, A. (2002). DNA methylation patterns and epigenetic memory. Genes & development *16*, 6-21.

Bird, A. P. (1980). DNA methylation and the frequency of CpG in animal DNA. Nucleic Acids Res *8*, 1499-1504.

Bird, A. P. (1986). CpG-rich islands and the function of DNA methylation. Nature *321*, 209-213.

Bird, A. P. (1993). Functions for DNA methylation in vertebrates. Cold Spring Harbor symposia on quantitative biology *58*, 281-285.

Bostick, M., Kim, J. K., Esteve, P. O., Clark, A., Pradhan, S., and Jacobsen, S. E. (2007). UHRF1 plays a role in maintaining DNA methylation in mammalian cells. Science *317*, 1760-1764.

Bracken, A. P., and Helin, K. (2009). Polycomb group proteins: navigators of lineage pathways led astray in cancer. Nat Rev Cancer *9*, 773-784.

Brahmer, J. R., Tykodi, S. S., Chow, L. Q., Hwu, W. J., Topalian, S. L., Hwu, P., Drake, C. G., Camacho, L. H., Kauh, J., Odunsi, K.*, et al.* (2012). Safety and activity of anti-PD-L1 antibody in patients with advanced cancer. N Engl J Med *366*, 2455-2465.

Brinkhuizen, T., van den Hurk, K., Winnepenninckx, V. J., de Hoon, J. P., van Marion, A. M., Veeck, J., van Engeland, M., and van Steensel, M. A. (2012). Epigenetic changes in Basal Cell Carcinoma affect SHH and WNT signaling components. PLoS ONE *7*, e51710.

Burdon, R. H., and Adams, R. L. (1969). The in vivo methylation of DNA in mouse fibroblasts. Biochimica et biophysica acta *174*, 322-329.

Campos, S. M., and Dizon, D. S. (2012). Antimitotic inhibitors. Hematol Oncol Clin North Am *26*, 607-628, viii-ix.

Cashen, A. F., Schiller, G. J., O'Donnell, M. R., and DiPersio, J. F. (2010). Multicenter, phase II study of decitabine for the first-line treatment of older patients with acute myeloid leukemia. Journal of clinical oncology : official journal of the American Society of Clinical Oncology *28*, 556-561.

Challa-Malladi, M., Lieu, Y. K., Califano, O., Holmes, A. B., Bhagat, G., Murty, V. V., Dominguez-Sola, D., Pasqualucci, L., and Dalla-Favera, R. (2011). Combined genetic inactivation of beta2-Microglobulin and CD58 reveals frequent escape from immune recognition in diffuse large B cell lymphoma. Cancer Cell *20*, 728-740.

Chatterjee, A., Mambo, E., Zhang, Y., Deweese, T., and Sidransky, D. (2006). Targeting of mutant hogg1 in mammalian mitochondria and nucleus: effect on cellular survival upon oxidative stress. BMC Cancer *6*, 235.

Chen, H. Y., Zhu, B. H., Zhang, C. H., Yang, D. J., Peng, J. J., Chen, J. H., Liu, F. K., and He, Y. L. (2012). High CpG island methylator phenotype is associated with lymph node metastasis and prognosis in gastric cancer. Cancer science *103*, 73-79.

Chen, T., Hevi, S., Gay, F., Tsujimoto, N., He, T., Zhang, B., Ueda, Y., and Li, E. (2007). Complete inactivation of DNMT1 leads to mitotic catastrophe in human cancer cells. Nat Genet *39*, 391-396.

Chen, T., Ueda, Y., Dodge, J. E., Wang, Z., and Li, E. (2003). Establishment and maintenance of genomic methylation patterns in mouse embryonic stem cells by Dnmt3a and Dnmt3b. Mol Cell Biol *23*, 5594-5605.

Cheng, X., and Blumenthal, R. M. (2008). Mammalian DNA methyltransferases: a structural perspective. Structure *16*, 341-350.

Cheriyath, V., Leaman, D. W., and Borden, E. C. (2011). Emerging roles of FAM14 family members (G1P3/ISG 6-16 and ISG12/IFI27) in innate immunity and cancer. J Interferon Cytokine Res *31*, 173-181.

Chodavarapu, R. K., Feng, S., Bernatavichute, Y. V., Chen, P. Y., Stroud, H., Yu, Y., Hetzel, J. A., Kuo, F., Kim, J., Cokus, S. J.*, et al.* (2010). Relationship between nucleosome positioning and DNA methylation. Nature *466*, 388-392.

Chow, C., Wong, N., Pagano, M., Lun, S. W., Nakayama, K. I., Nakayama, K., and Lo, K. W. (2012). Regulation of APC/CCdc20 activity by RASSF1A-APC/CCdc20 circuitry. Oncogene *31*, 1975-1987.

Chuang, L. S., Ian, H. I., Koh, T. W., Ng, H. H., Xu, G., and Li, B. F. (1997). Human DNA-(cytosine-5) methyltransferase-PCNA complex as a target for p21WAF1. Science *277*, 1996-2000.

Clapier, C. R., and Cairns, B. R. (2009). The biology of chromatin remodeling complexes. Annual review of biochemistry *78*, 273-304.

Claus, R., Almstedt, M., and Lubbert, M. (2005). Epigenetic treatment of hematopoietic malignancies: in vivo targets of demethylating agents. Semin Oncol *32*, 511-520.

Corona, D. F., and Tamkun, J. W. (2004). Multiple roles for ISWI in transcription, chromosome organization and DNA replication. Biochimica et biophysica acta *1677*, 113-119.

Dammann, R., Li, C., Yoon, J. H., Chin, P. L., Bates, S., and Pfeifer, G. P. (2000). Epigenetic inactivation of a RAS association domain family protein from the lung tumour suppressor locus 3p21.3. Nat Genet *25*, 315-319.

Davis, S., Du, P., Bilke, S., Triche, T. J., and Bootwalla, M. (2012). methylumi: Handle Illumina methylation data. R package version 2.2.0.

De Carvalho, D. D., Sharma, S., You, J. S., Su, S. F., Taberlay, P. C., Kelly, T. K., Yang, X., Liang, G., and Jones, P. A. (2012). DNA methylation screening identifies driver epigenetic events of cancer cell survival. Cancer Cell *21*, 655-667.

de Vos, D., and van Overveld, W. (2005). Decitabine: a historical review of the development of an epigenetic drug. Annals of hematology *84 Suppl 1*, 3-8.

Denslow, S. A., and Wade, P. A. (2007). The human Mi-2/NuRD complex and gene regulation. Oncogene *26*, 5433-5438.

Dimova, D. K., and Dyson, N. J. (2005). The E2F transcriptional network: old acquaintances with new faces. Oncogene *24*, 2810-2826.

Dunn, G. P., Bruce, A. T., Ikeda, H., Old, L. J., and Schreiber, R. D. (2002). Cancer immunoediting: from immunosurveillance to tumor escape. Nat Immunol *3*, 991-998.

Easwaran, H., Johnstone, S. E., Van Neste, L., Ohm, J., Mosbruger, T., Wang, Q., Aryee, M. J., Joyce, P., Ahuja, N., Weisenberger, D.*, et al.* (2012). A DNA hypermethylation module for the stem/progenitor cell signature of cancer. Genome research *22*, 837-849.

Easwaran, H. P., Van Neste, L., Cope, L., Sen, S., Mohammad, H. P., Pageau, G. J., Lawrence, J. B., Herman, J. G., Schuebel, K. E., and Baylin, S. B. (2010). Aberrant silencing of cancer-related genes by CpG hypermethylation occurs independently of their spatial organization in the nucleus. Cancer Res *70*, 8015-8024.

Egger, G., Jeong, S., Escobar, S. G., Cortez, C. C., Li, T. W., Saito, Y., Yoo, C. B., Jones, P. A., and Liang, G. (2006). Identification of DNMT1 (DNA methyltransferase 1) hypomorphs in somatic knockouts suggests an essential role for DNMT1 in cell survival. Proc Natl Acad Sci U S A *103*, 14080-14085.

Eissenberg, J. C., and Shilatifard, A. (2010). Histone H3 lysine 4 (H3K4) methylation in development and differentiation. Developmental biology *339*, 240-249.

Espada, J., Ballestar, E., Santoro, R., Fraga, M. F., Villar-Garea, A., Nemeth, A., Lopez-Serra, L., Ropero, S., Aranda, A., Orozco, H.*, et al.* (2007). Epigenetic disruption of ribosomal RNA genes and nucleolar architecture in DNA methyltransferase 1 (Dnmt1) deficient cells. Nucleic Acids Res *35*, 2191-2198. Epub 2007 Mar 2113.

Esteller, M., Corn, P. G., Baylin, S. B., and Herman, J. G. (2001). A gene hypermethylation profile of human cancer. Cancer Res *61*, 3225-3229.

Fadloun, A., Eid, A., and Torres-Padilla, M. E. (2013). Mechanisms and dynamics of heterochromatin formation during mammalian development: closed paths and open questions. Current topics in developmental biology *104*, 1-45.

Fahrner, J. A., and Baylin, S. B. (2003). Heterochromatin: stable and unstable invasions at home and abroad. Genes & development *17*, 1805-1812.

Fan, W., and Luo, J. (2010). SIRT1 regulates UV-induced DNA repair through deacetylating XPA. Mol Cell *39*, 247-258.

Federico, A., Morgillo, F., Tuccillo, C., Ciardiello, F., and Loguercio, C. (2007). Chronic inflammation and oxidative stress in human carcinogenesis. Int J Cancer *121*, 2381-2386.

Feinberg, A. P. (2007). Phenotypic plasticity and the epigenetics of human disease. Nature *447*, 433-440.

Felsenfeld, G., and McGhee, J. D. (1986). Structure of the 30 nm chromatin fiber. Cell *44*, 375-377.

Folle, G. A., Liddle, P., Lafon-Hughes, L., and Di Tomaso, M. V. (2010). Close encounters: RIDGEs, hyperacetylated chromatin, radiation breakpoints and genes differentially expressed in tumors cluster at specific human chromosome regions. Cytogenet Genome Res *128*, 17-27.

Fonsatti, E., Nicolay, H. J., Sigalotti, L., Calabro, L., Pezzani, L., Colizzi, F., Altomonte, M., Guidoboni, M., Marincola, F. M., and Maio, M. (2007). Functional up-regulation of human leukocyte antigen class I antigens expression by 5-aza-2'-deoxycytidine in cutaneous melanoma: immunotherapeutic implications. Clin Cancer Res *13*, 3333-3338.

Gabbara, S., and Bhagwat, A. S. (1995). The mechanism of inhibition of DNA (cytosine-5-)-methyltransferases by 5-azacytosine is likely to involve methyl transfer to the inhibitor. Biochem J *307 (Pt 1)*, 87-92.

Glasmacher, E., Agrawal, S., Chang, A. B., Murphy, T. L., Zeng, W., Vander Lugt, B., Khan, A. A., Ciofani, M., Spooner, C., Rutz, S.*, et al.* (2012). A Genomic Regulatory Element That Directs Assembly and Function of Immune-Specific AP-1-IRF Complexes. Science *338*, 975-80.

Goodwin, A. C., Shields, C. E., Wu, S., Huso, D. L., Wu, X., Murray-Stewart, T. R., Hacker-Prietz, A., Rabizadeh, S., Woster, P. M., Sears, C. L., and Casero, R. A., Jr. (2011). Polyamine catabolism contributes to enterotoxigenic Bacteroides fragilis-induced colon tumorigenesis. Proc Natl Acad Sci U S A *108*, 15354-15359.

Gore, S. D. (2005). Combination therapy with DNA methyltransferase inhibitors in hematologic malignancies. Nat Clin Pract Oncol *2 Suppl 1*, S30-35.

Goto, S., Bono, H., Ogata, H., Fujibuchi, W., Nishioka, T., Sato, K., and Kanehisa, M. (1997). Organizing and computing metabolic pathway data in terms of binary relations. Pacific Symposium on Biocomputing Pacific Symposium on Biocomputing, 175-186.

Graff, J., and Tsai, L. H. (2013). Histone acetylation: molecular mnemonics on the chromatin. Nature reviews Neuroscience *14*, 97-111.

Groselj, B., Sharma, N. L., Hamdy, F. C., Kerr, M., and Kiltie, A. E. (2013). Histone deacetylase inhibitors as radiosensitisers: effects on DNA damage signalling and repair. Br J Cancer *108*, 748-754.

Ha, K., Lee, G. E., Palii, S. S., Brown, K. D., Takeda, Y., Liu, K., Bhalla, K. N., and Robertson, K. D. (2010). Rapid and transient recruitment of DNMT1 to DNA double-strand breaks is mediated by its interaction with multiple components of the DNA damage response machinery. Hum Mol Genet *20*, 126-140.

Hahn, M. A., Hahn, T., Lee, D. H., Esworthy, R. S., Kim, B. W., Riggs, A. D., Chu, F. F., and Pfeifer, G. P. (2008). Methylation of polycomb target genes in intestinal cancer is mediated by inflammation. Cancer Res *68*, 10280-10289.

Hammerman, P. S., Hayes, D. N., Wilkerson, M. D., Schultz, N., Bose, R., Chu, A., Collisson, E. A., Cope, L., Creighton, C. J., Getz, G.*, et al.* (2012). Comprehensive genomic characterization of squamous cell lung cancers. Nature *489*, 519-525.

Hargreaves, D. C., and Crabtree, G. R. (2011). ATP-dependent chromatin remodeling: genetics, genomics and mechanisms. Cell research *21*, 396-420.

Herman, J. G., and Baylin, S. B. (2003). Gene silencing in cancer in association with promoter hypermethylation. N Engl J Med *349*, 2042-2054.

Hernandez, J. M., Floyd, D. H., Weilbaecher, K. N., Green, P. L., and Boris-Lawrie, K. (2008). Multiple facets of junD gene expression are atypical among AP-1 family members. Oncogene *27*, 4757-4767.

Hsu, T. H., Chu, C. C., Jiang, S. Y., Hung, M. W., Ni, W. C., Lin, H. E., and Chang, T. C. (2012). Expression of the class II tumor suppressor gene RIG1 is directly regulated by p53 tumor suppressor in cancer cell lines. FEBS Lett *586*, 1287-1293.

Ishikawa, H., Ma, Z., and Barber, G. N. (2009). STING regulates intracellular DNA-mediated, type I interferon-dependent innate immunity. Nature *461*, 788-792.

Issa, J. P. (2000). CpG-island methylation in aging and cancer. Current topics in microbiology and immunology *249*, 101-118.

Issa, J. P., Garcia-Manero, G., Giles, F. J., Mannari, R., Thomas, D., Faderl, S., Bayar, E., Lyons, J., Rosenfeld, C. S., Cortes, J., and Kantarjian, H. M. (2004). Phase 1 study of low-dose prolonged exposure schedules of the hypomethylating agent 5-aza-2'-deoxycytidine (decitabine) in hematopoietic malignancies. Blood *103*, 1635-1640.

Issa, J. P., and Kantarjian, H. (2005). Azacitidine. Nat Rev Drug Discov *Suppl*, S6-7.

Issa, J. P., and Kantarjian, H. M. (2009). Targeting DNA methylation. Clin Cancer Res *15*, 3938-3946.

Jaenisch, R., Harbers, K., Jahner, D., Stewart, C., and Stuhlmann, H. (1982). DNA methylation, retroviruses, and embryogenesis. Journal of cellular biochemistry *20*, 331-336.

Jee, C. D., Kim, M. A., Jung, E. J., Kim, J., and Kim, W. H. (2009). Identification of genes epigenetically silenced by CpG methylation in human gastric carcinoma. Eur J Cancer *45*, 1282-1293.

Jenuwein, T., and Allis, C. D. (2001). Translating the histone code. Science *293*, 1074-1080.

Jeong, S., Liang, G., Sharma, S., Lin, J. C., Choi, S. H., Han, H., Yoo, C. B., Egger, G., Yang, A. S., and Jones, P. A. (2009). Selective anchoring of DNA methyltransferases 3A and 3B to nucleosomes containing methylated DNA. Mol Cell Biol *29*, 5366-5376.

Johnson, J. A., Roden, D. M., Lesko, L. J., Ashley, E., Klein, T. E., and Shuldiner, A. R. (2012). Clopidogrel: a case for indication-specific pharmacogenetics. Clinical pharmacology and therapeutics *91*, 774-776.

Jones, P. A., and Baylin, S. B. (2007). The epigenomics of cancer. Cell *128*, 683-692.

Jones, P. A., and Liang, G. (2009). Rethinking how DNA methylation patterns are maintained. Nature reviews Genetics *10*, 805-811.

Jones, P. A., and Taylor, S. M. (1980). Cellular differentiation, cytidine analogs and DNA methylation. Cell *20*, 85-93.

Jones, P. A., and Taylor, S. M. (1981). Hemimethylated duplex DNAs prepared from 5-azacytidine-treated cells. Nucleic Acids Res *9*, 2933-2947.

Joshi-Tope, G., Gillespie, M., Vastrik, I., D'Eustachio, P., Schmidt, E., de Bono, B., Jassal, B., Gopinath, G. R., Wu, G. R., Matthews, L.*, et al.* (2005). Reactome: a knowledgebase of biological pathways. Nucleic Acids Res *33*, D428-432.

Juergens, R. A., Wrangle, J., Vendetti, F. P., Murphy, S. C., Zhao, M., Coleman, B., Sebree, R., Rodgers, K., Hooker, C. M., Franco, N.*, et al.* (2011). Combination epigenetic therapy has efficacy in patients with refractory advanced non-small cell lung cancer. Cancer Discov *1*, 598-607.

Kantarjian, H., Issa, J. P., Rosenfeld, C. S., Bennett, J. M., Albitar, M., DiPersio, J., Klimek, V., Slack, J., de Castro, C., Ravandi, F.*, et al.* (2006). Decitabine improves patient outcomes in myelodysplastic syndromes: results of a phase III randomized study. Cancer *106*, 1794-1803.

Kantarjian, H. M., O'Brien, S., Cortes, J., Giles, F. J., Faderl, S., Issa, J. P., Garcia-Manero, G., Rios, M. B., Shan, J., Andreeff, M.*, et al.* (2003). Results of decitabine (5-aza-2'deoxycytidine) therapy in 130 patients with chronic myelogenous leukemia. Cancer *98*, 522-528.

Keshet, I., Lieman-Hurwitz, J., and Cedar, H. (1986). DNA methylation affects the formation of active chromatin. Cell *44*, 535-543.

Khorasanizadeh, S. (2004). The nucleosome: from genomic organization to genomic regulation. Cell *116*, 259-272.

Kim, G. D., Ni, J., Kelesoglu, N., Roberts, R. J., and Pradhan, S. (2002). Co-operation and communication between the human maintenance and de novo DNA (cytosine-5) methyltransferases. Embo J *21*, 4183-4195.

Kim, S. Y., Levenson, J. M., Korsmeyer, S., Sweatt, J. D., and Schumacher, A. (2007). Developmental regulation of Eed complex composition governs a switch in global histone modification in brain. J Biol Chem *282*, 9962-9972.

Kulaeva, O. I., Draghici, S., Tang, L., Kraniak, J. M., Land, S. J., and Tainsky, M. A. (2003). Epigenetic silencing of multiple interferon pathway genes after cellular immortalization. Oncogene *22*, 4118-4127.

Kuzmichev, A., Margueron, R., Vaquero, A., Preissner, T. S., Scher, M., Kirmizis, A., Ouyang, X., Brockdorff, N., Abate-Shen, C., Farnham, P., and Reinberg, D. (2005). Composition and histone substrates of polycomb repressive group complexes change during cellular differentiation. Proc Natl Acad Sci U S A *102*, 1859-1864.

Lai, A. Y., and Wade, P. A. (2011). Cancer biology and NuRD: a multifaceted chromatin remodelling complex. Nat Rev Cancer *11*, 588-596.

Lamond, A. I., and Earnshaw, W. C. (1998). Structure and function in the nucleus. Science *280*, 547-553.

Leonhardt, H., Page, A. W., Weier, H. U., and Bestor, T. H. (1992). A targeting sequence directs DNA methyltransferase to sites of DNA replication in mammalian nuclei. Cell *71*, 865-873.

Li, E., Bestor, T. H., and Jaenisch, R. (1992). Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. Cell *69*, 915-926.

Li, J., Langst, G., and Grummt, I. (2006). NoRC-dependent nucleosome positioning silences rRNA genes. EMBO J *25*, 5735-5741.

Li, Q., and Tainsky, M. A. (2011). Epigenetic silencing of IRF7 and/or IRF5 in lung cancer cells leads to increased sensitivity to oncolytic viruses. PLoS One *6*, e28683.

Lilley, D. M., and Pardon, J. F. (1979). Structure and function of chromatin. Annual review of genetics *13*, 197-233.

Lister, R., Pelizzola, M., Dowen, R. H., Hawkins, R. D., Hon, G., Tonti-Filippini, J., Nery, J. R., Lee, L., Ye, Z., Ngo, Q. M.*, et al.* (2009). Human DNA methylomes at base resolution show widespread epigenomic differences. Nature *462*, 315-322.

Liu, J. B., Zhang, Y. X., Zhou, S. H., Shi, M. X., Cai, J., Liu, Y., Chen, K. P., and Qiang, F. L. (2011). CpG island methylator phenotype in plasma is associated with hepatocellular carcinoma prognosis. World journal of gastroenterology : WJG *17*, 4718-4724.

Logan, M. R., Jordan-Williams, K. L., Poston, S., Liao, J., and Taparowsky, E. J. (2012). Overexpression of Batf induces an apoptotic defect and an associated lymphoproliferative disorder in mice. Cell Death Dis *3*, e310.

Lu, R., Au, W. C., Yeow, W. S., Hageman, N., and Pitha, P. M. (2000). Regulation of the promoter activity of interferon regulatory factor-7 gene. Activation by interferon snd silencing by hypermethylation. J Biol Chem *275*, 31805-31812.

Lukas, J., Lukas, C., and Bartek, J. (2011). More than just a focus: The chromatin response to DNA damage and its role in genome integrity maintenance. Nat Cell Biol *13*, 1161-1169.

Marfella, C. G., and Imbalzano, A. N. (2007). The Chd family of chromatin remodelers. Mutation research *618*, 30-40.

Marks, H., Kalkan, T., Menafra, R., Denissov, S., Jones, K., Hofemeister, H., Nichols, J., Kranz, A., Stewart, A. F., Smith, A., and Stunnenberg, H. G. (2012). The transcriptional and epigenomic foundations of ground state pluripotency. Cell *149*, 590-604.

Marks, P., Rifkind, R. A., Richon, V. M., Breslow, R., Miller, T., and Kelly, W. K. (2001). Histone deacetylases and cancer: causes and therapies. Nat Rev Cancer *1*, 194-202.

McGarvey, K. M., Greene, E., Fahrner, J. A., Jenuwein, T., and Baylin, S. B. (2007). DNA methylation and complete transcriptional silencing of cancer genes persist after depletion of EZH2. Cancer Res *67*, 5097-5102.

McGarvey, K. M., Van Neste, L., Cope, L., Ohm, J. E., Herman, J. G., Van Criekinge, W., Schuebel, K. E., and Baylin, S. B. (2008). Defining a chromatin pattern that characterizes DNA-hypermethylated genes in colon cancer cells. Cancer Res *68*, 5753-5759.

McGhee, J. D., and Felsenfeld, G. (1980). Nucleosome structure. Annual review of biochemistry *49*, 1115-1156.

Mills, K. D., Sinclair, D. A., and Guarente, L. (1999). MEC1-dependent redistribution of the Sir3 silencing protein from telomeres to DNA double-strand breaks. Cell *97*, 609-620.

Min, J. N., Tian, Y., Xiao, Y., Wu, L., Li, L., and Chang, S. (2013). The mINO80 chromatin remodeling complex is required for efficient telomere replication and maintenance of genome stability. Cell research *23*, 1396-1413.

Mohrmann, L., and Verrijzer, C. P. (2005). Composition and functional specificity of SWI2/SNF2 class chromatin remodeling complexes. Biochimica et biophysica acta *1681*, 59-73.

Montero, J. C., Seoane, S., Ocana, A., and Pandiella, A. (2011). Inhibition of SRC family kinases and receptor tyrosine kinases by dasatinib: possible combinations in solid tumors. Clin Cancer Res *17*, 5546-5552.

Narlikar, G. J., Sundaramoorthy, R., and Owen-Hughes, T. (2013). Mechanisms and functions of ATP-dependent chromatin-remodeling enzymes. Cell *154*, 490-503.

O'Hagan, H. M., Mohammad, H. P., and Baylin, S. B. (2008). Double strand breaks can initiate gene silencing and SIRT1-dependent onset of DNA methylation in an exogenous promoter CpG island. PLoS Genet *4*, e1000155.

Oberdoerffer, P., Michan, S., McVay, M., Mostoslavsky, R., Vann, J., Park, S. K., Hartlerode, A., Stegmuller, J., Hafner, A., Loerch, P.*, et al.* (2008). SIRT1 redistribution on chromatin promotes genomic stability but alters gene expression during aging. Cell *135*, 907-918.

Ohm, J. E., McGarvey, K. M., Yu, X., Cheng, L., Schuebel, K. E., Cope, L., Mohammad, H. P., Chen, W., Daniel, V. C., Yu, W.*, et al.* (2007). A stem cell-like chromatin pattern may predispose tumor suppressor genes to DNA hypermethylation and heritable silencing. Nat Genet *39*, 237-242.

Okano, M., Bell, D. W., Haber, D. A., and Li, E. (1999). DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. Cell *99*, 247-257.

Okano, M., Xie, S., and Li, E. (1998). Cloning and characterization of a family of novel mammalian DNA (cytosine-5) methyltransferases. Nat Genet *19*, 219-220.

Oki, Y., Aoki, E., and Issa, J. P. (2007). Decitabine--bedside to bench. Critical reviews in oncology/hematology *61*, 140-152.

Ortega-Cava, C. F., Ishihara, S., Rumi, M. A., Aziz, M. M., Kazumori, H., Yuki, T., Mishima, Y., Moriyama, I., Kadota, C., Oshima, N.*, et al.* (2006). Epithelial toll-like receptor 5 is constitutively localized in the mouse cecum and exhibits distinctive down-regulation during experimental colitis. Clin Vaccine Immunol *13*, 132-138.

Pan, D. (2010). The hippo signaling pathway in development and cancer. Dev Cell *19*, 491-505.

Pardoll, D. M. (2012). The blockade of immune checkpoints in cancer immunotherapy. Nat Rev Cancer *12*, 252-264.

Pfeifer, G. P., Kadam, S., and Jin, S. G. (2013). 5-hydroxymethylcytosine and its potential roles in development and cancer. Epigenetics & chromatin *6*, 10.

Piccolo, F. M., and Fisher, A. G. (2013). Getting rid of DNA methylation. Trends in cell biology *2*,136-43.

Plass, C., Pfister, S. M., Lindroth, A. M., Bogatyrova, O., Claus, R., and Lichter, P. (2013). Mutations in regulators of the epigenome and their connections to global chromatin patterns in cancer. Nature reviews Genetics *14*, 765-780.

Price, B. D., and D'Andrea, A. D. (2013). Chromatin remodeling at DNA double-strand breaks. Cell *152*, 1344-1354.

Procko, E., and Gaudet, R. (2009). Antigen processing and presentation: TAPping into ABC transporters. Curr Opin Immunol *21*, 84-91.

Pruitt, K., Zinn, R. L., Ohm, J. E., McGarvey, K. M., Kang, S. H., Watkins, D. N., Herman, J. G., and Baylin, S. B. (2006). Inhibition of SIRT1 reactivates silenced cancer genes without loss of promoter DNA hypermethylation. PLoS Genet *2*, e40.

Raghavan, M., Del Cid, N., Rizvi, S. M., and Peters, L. R. (2008). MHC class I assembly: out and about. Trends Immunol *29*, 436-443.

Ravasi, T., Suzuki, H., Cannistraci, C. V., Katayama, S., Bajic, V. B., Tan, K., Akalin, A., Schmeier, S., Kanamori-Katayama, M., Bertin, N.*, et al.* (2010). An atlas of combinatorial transcriptional regulation in mouse and man. Cell *140*, 744-752.

Razin, A., and Cedar, H. (1977). Distribution of 5-methylcytosine in chromatin. Proc Natl Acad Sci U S A *74*, 2725-2728.

Reardon, J. T., Bessho, T., Kung, H. C., Bolton, P. H., and Sancar, A. (1997). In vitro repair of oxidative DNA damage by human nucleotide excision repair system: possible explanation for neurodegeneration in xeroderma pigmentosum patients. Proc Natl Acad Sci U S A *94*, 9463-9468.

Rhee, I., Bachman, K. E., Park, B. H., Jair, K. W., Yen, R. W., Schuebel, K. E., Cui, H., Feinberg, A. P., Lengauer, C., Kinzler, K. W.*, et al.* (2002). DNMT1 and DNMT3b cooperate to silence genes in human cancer cells. Nature *416*, 552-556.

Rhee, K. J., Wu, S., Wu, X., Huso, D. L., Karim, B., Franco, A. A., Rabizadeh, S., Golub, J. E., Mathews, L. E., Shin, J.*, et al.* (2009). Induction of persistent colitis by a human commensal, enterotoxigenic Bacteroides fragilis, in wild-type C57BL/6 mice. Infect Immun *77*, 1708-1718.

Robertson, K. D., Ait-Si-Ali, S., Yokochi, T., Wade, P. A., Jones, P. L., and Wolffe, A. P. (2000). DNMT1 forms a complex with Rb, E2F1 and HDAC1 and represses transcription from E2F-responsive promoters. Nat Genet *25*, 338-342.

Rogakou, E. P., Boon, C., Redon, C., and Bonner, W. M. (1999). Megabase chromatin domains involved in DNA double-strand breaks in vivo. J Cell Biol *146*, 905-916.

Rogakou, E. P., Pilch, D. R., Orr, A. H., Ivanova, V. S., and Bonner, W. M. (1998). DNA double-stranded breaks induce histone H2AX phosphorylation on serine 139. J Biol Chem *273*, 5858-5868.

Rountree, M. R., Bachman, K. E., and Baylin, S. B. (2000). DNMT1 binds HDAC2 and a new co-repressor, DMAP1, to form a complex at replication foci. Nat Genet *25*, 269-277.

Samarajiwa, S. A., Forster, S., Auchettl, K., and Hertzog, P. J. (2009). INTERFEROME: the database of interferon regulated genes. Nucleic Acids Res *37*, D852-857.

Santi, D. V., Norment, A., and Garrett, C. E. (1984). Covalent bond formation between a DNA-cytosine methyltransferase and DNA containing 5-azacytosine. Proc Natl Acad Sci U S A *81*, 6993-6997.

Santini, V., Kantarjian, H. M., and Issa, J. P. (2001). Changes in DNA methylation in neoplasia: pathophysiology and therapeutic implications. Annals of internal medicine *134*, 573-586.

Scarano, E., Iaccarino, M., Grippo, P., and Winckelmans, D. (1965). On methylation of DNA during development of the sea urchin embryo. Journal of molecular biology *14*, 603-607.

Schlesinger, Y., Straussman, R., Keshet, I., Farkash, S., Hecht, M., Zimmerman, J., Eden, E., Yakhini, Z., Ben-Shushan, E., Reubinoff, B. E.*, et al.* (2007). Polycomb-mediated methylation on Lys27 of histone H3 pre-marks genes for de novo methylation in cancer. Nat Genet *39*, 232-236.

Sharif, J., Muto, M., Takebayashi, S., Suetake, I., Iwamatsu, A., Endo, T. A., Shinga, J., Mizutani-Koseki, Y., Toyoda, T., Okamura, K.*, et al.* (2007). The SRA protein Np95 mediates epigenetic inheritance by recruiting Dnmt1 to methylated DNA. Nature *450*, 908-912.

Sharma, S., and Fitzgerald, K. A. (2010). Viral defense: it takes two MAVS to Tango. Cell *141*, 570-572.

Shen, H., and Laird, P. W. (2013). Interplay between the cancer genome and epigenome. Cell *153*, 38-55.

Shepherd, F. A., Rodrigues Pereira, J., Ciuleanu, T., Tan, E. H., Hirsh, V., Thongprasert, S., Campos, D., Maoleekoonpiroj, S., Smylie, M., Martins, R.*, et al.* (2005). Erlotinib in previously treated non-small-cell lung cancer. N Engl J Med *353*, 123-132.

Shinkai, Y., and Tachibana, M. (2011). H3K9 methyltransferase G9a and the related molecule GLP. Genes & development *25*, 781-788.

Siegel, R., Naishadham, D., and Jemal, A. (2013). Cancer statistics, 2013. CA Cancer J Clin *63*, 11-30.

Silverman, L. R., Demakos, E. P., Peterson, B. L., Kornblith, A. B., Holland, J. C., Odchimar-Reissig, R., Stone, R. M., Nelson, D., Powell, B. L., DeCastro, C. M.*, et al.* (2002). Randomized controlled trial of azacitidine in patients with the myelodysplastic syndrome: a study of the cancer and leukemia group B. Journal of clinical oncology : official journal of the American Society of Clinical Oncology *20*, 2429-2440.

Silverman, L. R., and Mufti, G. J. (2005). Methylation inhibitor therapy in the treatment of myelodysplastic syndrome. Nat Clin Pract Oncol *2 Suppl 1*, S12-23.

Simova, J., Pollakova, V., Indrova, M., Mikyskova, R., Bieblova, J., Stepanek, I., Bubenik, J., and Reinis, M. (2011). Immunotherapy augments the effect of 5-azacytidine on HPV16-associated tumours with different MHC class I-expression status. Br J Cancer *105*, 1533-1541.

Smyth, G. K., and Speed, T. (2003). Normalization of cDNA microarray data. Methods *31*, 265-273.

Solage, A., and Cedar, H. (1978). Organization of 5-methylcytosine in chromosomal DNA. Biochemistry *17*, 2934-2938.

Solomon, J. M., Pasupuleti, R., Xu, L., McDonagh, T., Curtis, R., DiStefano, P. S., and Huber, L. J. (2006). Inhibition of SIRT1 catalytic activity increases p53 acetylation but does not alter cell survival following DNA damage. Mol Cell Biol *26*, 28-38.

Song, M. S., Song, S. J., Ayad, N. G., Chang, J. S., Lee, J. H., Hong, H. K., Lee, H., Choi, N., Kim, J., Kim, H.*, et al.* (2004). The tumour suppressor RASSF1A regulates mitosis by inhibiting the APC-Cdc20 complex. Nat Cell Biol *6*, 129-137.

Soria, G., Polo, S. E., and Almouzni, G. (2012). Prime, repair, restore: the active role of chromatin in the DNA damage response. Mol Cell *46*, 722-734.

Spada, F., Haemmer, A., Kuch, D., Rothbauer, U., Schermelleh, L., Kremmer, E., Carell, T., Langst, G., and Leonhardt, H. (2007). DNMT1 but not its interaction with the replication machinery is required for maintenance of DNA methylation in human cells. J Cell Biol *176*, 565-571.

Spivakov, M., and Fisher, A. G. (2007). Epigenetic signatures of stem-cell identity. Nature reviews Genetics *8*, 263-271.

Steenken, S. (1997). Electron transfer in DNA? Competition by ultra-fast proton transfer? Biol Chem *378*, 1293-1297.

Strahl, B. D., and Allis, C. D. (2000). The language of covalent histone modifications. Nature *403*, 41-45.

Stresemann, C., and Lyko, F. (2008). Modes of action of the DNA methyltransferase inhibitors azacytidine and decitabine. Int J Cancer *123*, 8-13.

Strohner, R., Nemeth, A., Nightingale, K. P., Grummt, I., Becker, P. B., and Langst, G. (2004). Recruitment of the nucleolar remodeling complex NoRC establishes ribosomal DNA silencing in chromatin. Mol Cell Biol *24*, 1791-1798.

Strowig, T., Henao-Mejia, J., Elinav, E., and Flavell, R. (2012). Inflammasomes in health and disease. Nature *481*, 278-286.

Struhl, K., and Segal, E. (2013). Determinants of nucleosome positioning. Nature structural & molecular biology *20*, 267-273.

Subramanian, A., Tamayo, P., Mootha, V. K., Mukherjee, S., Ebert, B. L., Gillette, M. A., Paulovich, A., Pomeroy, S. L., Golub, T. R., Lander, E. S., and Mesirov, J. P. (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A *102*, 15545-15550.

Sugiyama, T., Cam, H. P., Sugiyama, R., Noma, K., Zofall, M., Kobayashi, R., and Grewal, S. I. (2007). SHREC, an effector complex for heterochromatic transcriptional silencing. Cell *128*, 491-504.

Tahiliani, M., Koh, K. P., Shen, Y., Pastor, W. A., Bandukwala, H., Brudno, Y., Agarwal, S., Iyer, L. M., Liu, D. R., Aravind, L., and Rao, A. (2009). Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. Science *324*, 930-935.

Takai, D., and Jones, P. A. (2002). Comprehensive analysis of CpG islands in human chromosomes 21 and 22. Proc Natl Acad Sci U S A *99*, 3740-3745.

Tamburini, B. A., and Tyler, J. K. (2005). Localized histone acetylation and deacetylation triggered by the homologous recombination pathway of double-strand DNA repair. Mol Cell Biol *25*, 4903-4913.

Teodoridis, J. M., Hardie, C., and Brown, R. (2008). CpG island methylator phenotype (CIMP) in cancer: causes and implications. Cancer letters *268*, 177-186.

Thomson, J. P., Skene, P. J., Selfridge, J., Clouaire, T., Guy, J., Webb, S., Kerr, A. R., Deaton, A., Andrews, R., James, K. D.*, et al.* (2010). CpG islands influence chromatin structure via the CpG-binding protein Cfp1. Nature *464*, 1082-1086.

Toedling, J., Skylar, O., Krueger, T., Fischer, J. J., Sperling, S., and Huber, W. (2007). Ringo--an R/Bioconductor package for analyzing ChIP-chip readouts. BMC Bioinformatics *8*, 221.

Topalian, S. L., Hodi, F. S., Brahmer, J. R., Gettinger, S. N., Smith, D. C., McDermott, D. F., Powderly, J. D., Carvajal, R. D., Sosman, J. A., Atkins, M. B.*, et al.* (2012). Safety, activity, and immune correlates of anti-PD-1 antibody in cancer. N Engl J Med *366*, 2443-2454.

Toyota, M., Ahuja, N., Ohe-Toyota, M., Herman, J. G., Baylin, S. B., and Issa, J. P. (1999). CpG island methylator phenotype in colorectal cancer. Proc Natl Acad Sci U S A *96*, 8681-8686.

Tsai, H. C., Li, H., Van Neste, L., Cai, Y., Robert, C., Rassool, F. V., Shin, J. J., Harbom, K. M., Beaty, R., Pappou, E.*, et al.* (2012). Transient Low Doses of DNA-Demethylating Agents Exert Durable Antitumor Effects on Hematological and Epithelial Tumor Cells. Cancer Cell *21*, 430-446.

Turcan, S., Rohle, D., Goenka, A., Walsh, L. A., Fang, F., Yilmaz, E., Campos, C., Fabius, A. W., Lu, C., Ward, P. S.*, et al.* (2012). IDH1 mutation is sufficient to establish the glioma hypermethylator phenotype. Nature *483*, 479-483.

Vadakara, J., and Borghaei, H. (2012). Personalized medicine and treatment approaches in non-small-cell lung carcinoma. Pharmgenomics Pers Med *5*, 113-123.

Van der Velden, J., Janssen-Heininger, Y. M., Mandalapu, S., Scheller, E. V., Kolls, J. K., and Alcorn, J. F. (2012). Differential requirement for c-Jun N-terminal kinase 1 in lung inflammation and host defense. PLoS One *7*, e34638.

Vaquerizas, J. M., Kummerfeld, S. K., Teichmann, S. A., and Luscombe, N. M. (2009). A census of human transcription factors: function, expression and evolution. Nature reviews Genetics *10*, 252-263.

Varley, K. E., Gertz, J., Bowling, K. M., Parker, S. L., Reddy, T. E., Pauli-Behn, F., Cross, M. K., Williams, B. A., Stamatoyannopoulos, J. A., Crawford, G. E.*, et al.* (2013). Dynamic DNA methylation across diverse human cell lines and tissues. Genome research *23*, 555-567.

Venter, J. C., Adams, M. D., Myers, E. W., Li, P. W., Mural, R. J., Sutton, G. G., Smith, H. O., Yandell, M., Evans, C. A., Holt, R. A.*, et al.* (2001). The sequence of the human genome. Science *291*, 1304-1351.

Vire, E., Brenner, C., Deplus, R., Blanchon, L., Fraga, M., Didelot, C., Morey, L., Van Eynde, A., Bernard, D., Vanderwinden, J. M.*, et al.* (2006). The Polycomb group protein EZH2 directly controls DNA methylation. Nature *439*, 871-874.

Waddington, C. H. (2012). The epigenotype. 1942. International journal of epidemiology *41*, 10-13.

Wagner, E. J., and Carpenter, P. B. (2012). Understanding the language of Lys36 methylation at histone H3. Nature reviews Molecular cell biology *13*, 115-126.

Wagner, T., and Jung, M. (2012). New lysine methyltransferase drug targets in cancer. Nature biotechnology *30*, 622-623.

Wang, Y., Fischle, W., Cheung, W., Jacobs, S., Khorasanizadeh, S., and Allis, C. D. (2004). Beyond the double helix: writing and reading the histone code. Novartis Foundation symposium *259*, 3-17; discussion 17-21, 163-169.

Warnes, G. R. (2012). gplots: Various R programming tools for plotting data. R package version 2.11.0. .

Widschwendter, M., Fiegl, H., Egle, D., Mueller-Holzner, E., Spizzo, G., Marth, C., Weisenberger, D. J., Campan, M., Young, J., Jacobs, I., and Laird, P. W. (2007). Epigenetic stem cell signature in cancer. Nat Genet *39*, 157-158.

Wijermans, P., Lubbert, M., Verhoef, G., Bosly, A., Ravoet, C., Andre, M., and Ferrant, A. (2000). Low-dose 5-aza-2'-deoxycytidine, a DNA hypomethylating agent, for the treatment of high-risk myelodysplastic syndrome: a multicenter phase II study in elderly patients. Journal of clinical oncology : official journal of the American Society of Clinical Oncology *18*, 956-962.

Wilson, G. G., and Murray, N. E. (1991). Restriction and modification systems. Annual review of genetics *25*, 585-627.

Wu, S., Rhee, K. J., Albesiano, E., Rabizadeh, S., Wu, X., Yen, H. R., Huso, D. L., Brancati, F. L., Wick, E., McAllister, F.*, et al.* (2009). A human colonic commensal promotes colon tumorigenesis via activation of T helper type 17 T cell responses. Nat Med *15*, 1016-1022.

Wutz, A. (2011). Gene silencing in X-chromosome inactivation: advances in understanding facultative heterochromatin formation. Nature reviews Genetics *12*, 542-553.

Xiao, A., Li, H., Shechter, D., Ahn, S. H., Fabrizio, L. A., Erdjument-Bromage, H., Ishibe-Murakami, S., Wang, B., Tempst, P., Hofmann, K.*, et al.* (2009). WSTF regulates the H2A.X DNA damage response via a novel tyrosine kinase activity. Nature *457*, 57-62.

Xu, Y., Zhong, H., and Shi, W. (2010). MAVS protects cells from apoptosis by negatively regulating VDAC1. Mol Cell Biochem *375*, 219.

Yang, X., Dutta, U., and Shaw, L. M. (2010). SHP2 mediates the localized activation of Fyn downstream of the alpha6beta4 integrin to promote carcinoma invasion. Mol Cell Biol *30*, 5306-5317.

Yang, Z., and Klionsky, D. J. (2010). Mammalian autophagy: core molecular machinery and signaling regulation. Curr Opin Cell Biol *22*, 124-131.

Yildirim, O., Li, R., Hung, J. H., Chen, P. B., Dong, X., Ee, L. S., Weng, Z., Rando, O. J., and Fazzio, T. G. (2011). Mbd3/NURD complex regulates expression of 5-hydroxymethylcytosine marked genes in embryonic stem cells. Cell *147*, 1498-1510.

Youlden, D. R., Cramb, S. M., and Baade, P. D. (2008). The International Epidemiology of Lung Cancer: geographical distribution and secular trends. J Thorac Oncol *3*, 819-831.

Yuan, Z., Zhang, X., Sengupta, N., Lane, W. S., and Seto, E. (2007). SIRT1 regulates the function of the Nijmegen breakage syndrome protein. Mol Cell *27*, 149-162.

Zambelli, F., Pesole, G., and Pavesi, G. (2009). Pscan: finding over-represented transcription factor binding site motifs in sequences from co-regulated or co-expressed genes. Nucleic Acids Res *37*, W247-252.

Zentner, G. E., and Henikoff, S. (2013). Regulation of nucleosome dynamics by histone modifications. Nature structural & molecular biology *20*, 259-266.

Zhang, Y., and Reinberg, D. (2001). Transcription regulation by histone methylation: interplay between different covalent modifications of the core histone tails. Genes & development *15*, 2343-2360.

Curriculum Vitae

Wei Wang, Ph.D. Candidate
220 E. 70th St., Apt. 8D
New York, NY 10021
(410) 428.6353
wwang34@jhmi.edu
biowendy@gmail.com

EDUCATION

2005-present	**Johns Hopkins University School of Medicine**, Baltimore, MD Ph.D. candidate, Human Genetics Program **Johns Hopkins University School of Public Health**, Baltimore, MD Concurrent M.H.S. candidate, Department of Biostatistics
2003-2005	**University of Wisconsin-Madison**, Madison, WI M.S., Department of Genetics
1999-2003	**Tsinghua University**, Beijing, China B.S., Department of Biological Sciences and Biotechnology

RESEARCH EXPERIENCE

2006-present

Graduate Student: Johns Hopkins University School of Medicine, Human Genetics Program
Concurrent Master Student: Johns Hopkins University School of Public Health, Department of Biostatistics
RESEARCH: Understanding the role for DNA methyltransferases (DNMTs) and components of Polycomb Repressive Complex 4 (PRC4) in epigenetic gene regulation during oxidative damage and their potential effect on tumorigenesis.
Advisor: Dr. Stephen B. Baylin
RESEARCH: Analyzing the genomic and epigenomic alterations in non-small cell lung cancer (NSCLC) by transient low-dose Azacytidine.
Advisor: Dr. Stephen B. Baylin
RESEARCH: Analysis of Illumina Infinium Human Methylation 450K Beadchip data for 135 NSCLC patients.
Advisor: Dr. Stephen B. Baylin, Dr. Hongkai Ji

2004-2005 **Master Student**: University of Wisconsin-Madison, Department of Genetics
RESEARCH: Study the mechanism of nonsense mediated mRNA decay (NMD) using *C. elegans* as a model organism.
Advisor: Dr. Philip Anderson

2002-2003 **Undergraduate Research**: Tsinghua University, Department of Biological Sciences and Biotechnology
RESEARCH: Expression, purification and crystal structural analysis of human hBub protein.
Advisor: Dr. Zihe Rao

ABSTRACTS

John Wrangle, **Wei Wang**, Alexander Koch, Hariharan Easwaran, Helai Mohammad, Princy Parsana, Frank Vendetti, Kristen Rodgers, Xiaoyu Pan, Kirsten Harbom, Cynthia Zahnow, Janis Taube, Julie Brahmer, Peter Jones, Suzanne Topalian, Charles Rudin, Malcolm Brock, Drew Pardoll, Stephen Baylin. Epigenetic therapy and sensitization of lung cancer to immunotherapy. AACR Conference: Washington, D.C., 2013.

Wei Wang, Heather M. O'Hagan, Subhojit Sen, Stella S. Lee, Eriko G. Clements, Yi Cai, Leander Van Neste, Hariharan Easwaran, Stephen B. Baylin. Oxidative Damage Targets Complexes Containing DNA Methyltransferases, SIRT1 and Polycomb Members to Promoter CpG Islands. JHU GSA poster session, 2011.

Heather M. O'Hagan, **Wei Wang**, Subhojit Sen, Stella S. Lee, Eriko G. Clements, Yi Cai, Leander Van Neste, Hariharan Easwaran, Stephen B. Baylin. Oxidative Damage Targets Complexes Containing DNA Methyltransferases, SIRT1 and Polycomb Members to Promoter CpG Islands. AACR Conference: Orlando, 2011.

PUBLICATIONS

Wrangle J*, **Wang W***, Koch A*, Easwaran H, Mohammad HP, Vendetti F, Vancriekinge W, Demeyer T, Du Z, Parsana P, Rodgers K, Yen RW, Zahnow CA, Taube JM, Brahmer JR, Tykodi SS, Easton K, Carvajal RD, Jones PA, Laird PW, Weisenberger DJ, Tsai S, Juergens RA, Topalian SL, Rudin CM, Brock MV, Pardoll D, Baylin SB. Alterations of immune response of non-small cell lung cancer with Azacytidine. *Oncotarget.* 2013 Nov;4(11):2067-79.
** These authors contributed equally to this work*

Di Cello F, Flowers VL, Li H, Vecchio-Pagán B, Gordon B, Harbom K, Shin J, Beaty R, **Wang W**, Brayton C, Baylin SB, Zahnow CA. Cigarette smoke induces epithelial to mesenchymal transition and increases the metastatic ability of breast cancer cells. *Mol Cancer.* 2013 Aug 6;12:90.

Di Cello F, Cope L, Li H, Jeschke J, **Wang W**, Baylin SB, Zahnow CA. Methylation of the claudin 1 promoter is associated with loss of expression in estrogen receptor positive breast cancer. *PLoS One.* 2013 Jul 3;8(7):e68630.

Cai Y, Geutjes EJ, de Lint K, Roepman P, Bruurs L, Yu LR, **Wang W**, van Blijswijk J, Mohammad H, de Rink I, Bernards R, Baylin SB. The NuRD complex cooperates with DNMTs to maintain silencing of key colorectal tumor suppressor genes. *Oncogene.* 2013 May 27.

Jin L, Hanigan CL, Wu Y, **Wang W**, Park BH, Woster PM, Casero RA. Loss of LSD1 (lysine-specific demethylase 1) suppresses growth and alters gene expression of human colon cancer cells in a p53- and DNMT1(DNA methyltransferase 1)-independent manner. *Biochem J.* 2013 Jan 15;449(2):459-68.

O'Hagan HM*, **Wang W***, Sen S, Destefano Shields C, Lee SS, Zhang YW, Clements EG, Cai Y, Van Neste L, Easwaran H, Casero RA, Sears CL, Baylin SB. Oxidative damage targets complexes containing DNA methyltransferases, SIRT1, and polycomb members to promoter CpG Islands. *Cancer Cell.* 2011 Nov 15;20(5):606-19.
** These authors contributed equally to this work*

TEACHING EXPERIENCE

2006	Teaching assistant for graduate school course of "Genetics", Johns Hopkins University
2004	Teaching assistant for undergraduate course of "Genetics", University of Wisconsin-Madison

HONORS

Shun-De Wu couple foundation scholarship for academic performance, 2002
Commercial bank of Beijing foundation scholarship for academic performance, 2001
Tsinghua University freshman scholarship, 2000

Intended to be blank